

Sharing real and fictional reference*

Hans Kamp
University of Stuttgart
University of Texas, Austin

August 12, 2021

Abstract

This chapter extends the framework of MSDRT (Mental State Discourse Representation Theory) to the problem of reference in fiction, and to the role and function of fictional names. Central to the investigation is the notion of an Entity Representation (ER), a central feature of MSDRT and used previously in the communication-theoretic analysis of the pragmatics and semantics of non-fictional names in Kamp (2015). As argued in that paper, the use of proper names within a speech community leads to networks of connected ERs in the mental states of their users. These networks provide the names with a kind of intersubjective identity. In this respect fictional names resemble non-fictional names – those that refer to real entities, which exist in the actual world in which we live. The analysis of fictional names and fictional reference I will propose capitalizes on this resemblance.

1 Introduction

One of the signal achievements in the philosophy of the 20-th century was the battery of Kripke’s observations and arguments showing that names do

*Thanks to the audiences of the two workshops (Groningen, Sept. 2018, Uppsala, July 2019) which led to the present collection for their helpful suggestions. Very special thanks go to Emar Maier, who reduced a much longer manuscript, in which more was said about the role of ERs in communication about real referents and much that wasn’t immediately relevant to reference in fiction, to the more streamlined and condensed version you find here.

not refer via descriptive contents that their users may associate with them. A name becomes the name of its referent through some act of association – a ‘baptism’, as Kripke calls it – and from then on those present at the baptism can use the name and as they do so pass it on to other members of the speech community. In this way the name can spread within the community, reaching ever more of its members. As argued in Kamp (2015), the transmission of the command of a name N from one member to another can take different forms; but common to all those different forms is a kind of matching of Entity Representations (ERs), the ER implicated in the source’s (the speaker’s or author’s) use of N and the ER that the target (the listener or reader) employs to interpret this use of N .

More precisely, two central assumptions of Kamp (2015) are that (i) when a source uses a name N as part of some utterance, that always involves her use of an ER which for her represents the entity to which she wants to refer and (ii) the interpretation of her use of N by the recipient involves an ER on his part, which he takes to be coreferential with the ER used by the source. Such acts of use and interpretation establish links between the two ERs involved. And that is how networks of matching ERs come about and develop over time, reaching ever larger portions of the community.

Most of this works for fictional names as it does for names of real things. There is a big difference, of course, at the outset. The creation of a work of fiction (or a world of fiction, like those of Greek mythology or Egyptian mythology) is ipso facto the creation of its characters. And normally the characters are given names. But once a character has been introduced and given its name, the use that is then made of the name by those who are acquainted with the work or world of fiction and want to talk about it closely resembles our use of non-fictional names, which came into existence through an act of baptism.

How similar the ways are in which we handle non-fictional and fictional names can be seen when you look at names of which it is or has been unclear whether they are fictional or non-fictional. Perhaps the most notorious example of this in the philosophical literature is the biblical figure Jonah. Kripke discussed this example to show that what descriptive information users associate with a name need not have anything to do with what the name refers to. In the case of Jonah these were the things that are said about Jonah in the bible and since some of those are so obviously impossible (like being swallowed by a whale and then walking off unharmed after regurgitation), many people must have thought that Jonah was a fiction. But presumably not ev-

everybody drew this conclusion, some believing firmly that he was a historical figure all the same, while others were agnostic or in doubt. In the meantime biblical scholarship has established beyond reasonable doubt that there actually was a person to whom the bible refers as Jonah, but that he had none of the salient properties the bible ascribes to him. The point I want to draw attention to is that even before this conclusion was reached, using ‘Jonah’ to speak of the biblical Jonah was quite unproblematic, even between those who believed he was real, those who thought he must be legend and those who weren’t sure. Other examples from our cultural heritage are *Gilgamesh* and *King Arthur* – names that have been in the mouths and minds of countless speakers past and present, with diverse and changing opinions as to whether or not they were the names of historical figures. And for yet another example of the same: when father and three year old are engaged in talk about Santa, their understanding of what is being said will be very different – for one it is talk about the real world and for the other it is make-belief. And yet the ways in which they each operate with the name *Santa*, in speaking and in making sense of the words of the other, do not seem to be all that different. One of the main aims of this chapter is to articulate in what sense they aren’t all that different.

The close similarities in our use of fictional and non-fictional names are a special case of the similarities in the production and interpretation of fictional and non-fictional discourse generally. What sets fictional and non-fictional discourse apart is the significance of the communicated contents. But there is no real difference in the ways in which fictional and non-fictional content are linguistically encoded and decoded; nor is there, according to MSDRT (Mental State Discourse Representation Theory)¹, the formal framework that

¹MSDRT is an extension of DRT (Discourse Representation Theory), an approach to natural language semantics in which natural language expressions (including sentences and multi-sentence texts) are assigned so-called DRSs (Discourse representation Structures) as their logical forms. The DRSs are formulas of certain so-called *DRS languages*, independently defined formal languages with a syntax and model-theoretic semantics (specified along the lines of the currently standard presentations of formalisms like the Predicate Calculus or the Typed Lambda Calculus). In DRT the semantics of a natural language expression E is given by the model-theoretic semantics for the DRS $K(E)$ that the theory assigns to E (Kamp and Reyle, 1993). MSDRT extends DRT in that it includes DRSs which contain representations of mental states of cognitive agents. These mental state representations are composed of representations of propositional attitudes and ERs. The ERs that, according to MSDRT, are among the constituents of mental states will be the main players in this chapter. For formal details, see e.g. Kamp (2003); Kamp et al. (2011);

will be used in this chapter, between the ways in which fictional and non-fictional contents are mentally represented. Trying to capture the essence of these similarities has been one of the central aims of DRT from the start.

I will review those features of MSDRT that are directly relevant to the story about fictional names that this chapter will gradually make its way towards. Section 2 presents MSDRT’s general approach to the representation of mental states. Section 3 is devoted to the notion that will be central to all we have to say about reference and about names, that of an Entity Representation, and more particularly the Named Entity Representations that agents use to represent entities that they know by name. Section 4 is devoted to the semantics and pragmatics of the reception of fiction and our handling of fictional names as one aspect of that.

2 Representing attitudes in MSDRT

One of the central motivations for the development of MSDRT was the observation that ‘attitude reports’ can take a great variety of forms, and that more often than not they involve stretches consisting of several connected attributions, presenting an interconnected web of different but related attitudes.

Here is an example. Suppose I believe that there is a wonder drug for the mortal disease with which I have been diagnosed – cancer of the pancreas, say. If I have this belief (in addition to the belief that I have the disease), then in all likelihood I will also have the desire to obtain this drug. Suppose also that the belief about the drug is based on hearsay (or on the irresponsible announcements of some pharmaceutical company); in actual fact there is no such drug. But that won’t alter the fact that my attitudes concerning it – the belief that there is this drug and my desire to obtain it – are just as real as they would have been if the belief had been true. And that includes their ‘coreferentiality’: that belief and desire are ‘about the same thing’, even if there is no actual thing that they are both about.

We need a systematic, formally precise way of describing such attitude complexes, which can capture this kind of ‘mind-internal coreference’ as well as the contents of the individual attitudes of which they are made up. The

Maier (2015). The examples discussed in the course of the present chapter should provide enough of an informal grasp for understanding the main points that it is concerned to make about reference in fiction and non-fiction.

descriptions that MSDRT makes available for this purpose are illustrated by (1), as representation for the attitudinal complex described above, consisting of the two beliefs – that one has cancer of the pancreas and that there is an effective drug against it – together with the desire to obtain the drug.²

²In DRT it has long been customary to treat sentences and clauses as descriptions of states or events. Part of this treatment is that in the semantic representations of the describing sentences and clauses the described states and events are specified as being of the types that the sentences or clauses describe them as. The notation used for these type specifications puts the discourse referent for the state or event up front, followed by ‘:’ and then by the typing information. Often, as in the two examples from (1) just mentioned, the typing takes the form of the ‘core predication’ of the sentence or clause, consisting of the verb and its core arguments.

(i) e and s are drefs for eventualities, more specifically events and states, respectively. ‘ e : obtain’(i, d)’ means that e is an event of i obtaining d . (ii) ‘ $\subseteq$ ’ and ‘ $\prec$ ’ denote the relations of temporal inclusion and temporal precedence, and ‘ n ’ represents a time (here: the time at which the represented belief is being entertained). ‘ $n \subseteq s$ ’ expresses that the time represented by ‘ n ’ is included in the duration of the state represented by ‘ s ’. (iii) Words with primes, as in ‘drug’(d)’, are expressions in the metalanguage in which a semantics is given for some natural language fragment that includes the unprimed word, following the practice widely used in Montague Grammar and approaches building on it, according to which there is for each content word α of the natural language fragment a corresponding expression α' (of the same logical type as α) in the *metalanguage* (the language *in* which the semantics is given). In DRT and MSDRT these expressions are also treated as parts of the vocabulary of the DRS languages used to represent the natural language fragments containing the corresponding unprimed words.

I have made an effort to be as explicit as possible about notation and background assumptions, but constraints on length for contributions to the present collection like this one had to be taken into account as well. (Even as it is, I can be accused of having overstretched these constraints; I’d like to thank the editors for their tolerance on this point.)

$$(1) \left\{ \begin{array}{l} \left\langle BEL, \begin{array}{|c|} \hline s \quad x \\ \hline n \subseteq s \quad \text{cancer.o.t.pancreas}'(x) \\ s: \text{have}'(i,x) \end{array} \right\rangle \\ \left\langle BEL, \begin{array}{|c|} \hline s' \quad d \\ \hline n \subseteq s' \quad \text{drug}'(d) \\ s': \text{cure-for}'(d,x) \end{array} \right\rangle \\ \left\langle DES, \begin{array}{|c|} \hline e \\ \hline n \prec e \\ e: \text{obtain}'(i,d) \end{array} \right\rangle \end{array} \right\}$$

Before we add entity representations to structures like (1) in Section 3 below, it may be useful to highlight two characteristics of ADF that are illustrated in (1).³

³For a little more background: MSDRT builds on DRT in two ways. First, it provides representation languages whose representations are structures like (1), which serve as (linguistically and cognitively relevant) representations of mental states. In this chapter I will sometimes refer to this first level of MSDRT as *ADF* (for *Attitude Description Formalism*). (The name ADF also records the close affinity between MSDRT and Maier's (2016) project, which uses the term Attitude Description Theory (ADT).) The second level of MSDRT serves to account for the semantics of attitude reports, including reports involving two or more distinct propositional attitudes with referentially connected contents, as are found in a sentence like (i).

- (i) John believes that there is a drug that cures cancer of the pancreas and he wants to get hold of it.

In its account of the semantics of attitude reports like (i) this level of MSDRT makes use of DRS-like representations that contain structures like (1) as constituents. The representation languages which include these representations among their DRSs are extensions of DRT with a model theory whose models contain (among many other types of entities) agents with mental states that are structured along the lines revealed by ADF structures like (1). The models for this full-fledged version of MSDRT articulate, as part of the truth conditions they provide for MSDRT's DRSs, what it means for a structure like (1) to be a correct description of someone's mental state.

The second level of MSDRT will not be needed in this chapter. So there is no need to go into its details here. I will also forego a formal definition of the representation language of ADF that is presupposed here (and that encompasses all the ADF structures which will

The first has to do with the special purpose dref n . The occurrence of n in the content representation of the first belief in (1) indicates that the content of this belief is that the agent (represented by the special dref i) has cancer of the pancreas *at the time when the agent is in the mental state depicted by (1)*.

The second feature characteristic of ADF arises when two or more attitude content DRSs are part of the same mental state description and share a discourse referent (or ‘dref’, the abbreviation I will often use from now on), in the way that the second belief content DRS $K_{BEL,2}$ and the desire content DRS K_{DES} in (1) share the dref d . In (1) this dref is ‘declared’ in the Universe of $K_{BEL,2}$, and because of this the content of this DRS is to the effect that there exists a drug d which cures cancer of the pancreas, just as in standard DRT. But d is then reused in K_{DES} , making the content of this DRS dependent on that of the belief DRS. In the model theory for ADF this dependence is made explicit by treating the semantic content of the desire DRS as an *information update* in the sense of Dynamic Semantics – i.e. as a *context change potential* or *CCP*. CCPs are functions from given *information states* to other, more informative, information states. For the combination of the second belief and the desire in (1) this gives an interpretation consisting of (i) the belief that there is a drug which cures cancer of the pancreas and (ii) building on this belief, a desire to obtain this drug.

There is also internal coreference between the first and the second belief in (1), mediated by the dref x . This case is less of a challenge, at least so long as we make the plausible assumption that there actually exists a disease by this name and that the agent knows this. We return to this in the next section.

3 Entity Representations

3.1 Representing Entities as distinct from representing Propositional Content

From this presentation of MSDRT one crucial ingredient is still missing. According to MSDRT mental states are made up not just of propositional

be used in what follows). Explanations of notation used in these ADF structures will be provided as we go along. Readers who want more details about ADF or about full MSDRT are referred to the publications mentioned in footnote 1.

attitudes, but of combinations of propositional attitudes and entity representations. This assumption isn't just one of theoretical opportunism, which proves to be convenient when MSDRT is used as logical form formalism for the representation of simple and complex attitude reports. Intuitively it seems plain that the mental state of pretty much everyone of us is loaded with entity representations; our mental states include whole libraries of them, representations of all the different kinds of entities that we have encountered or learned about in the course of our lives and that are still with us, in that they haven't been expunged from our memories. Moreover, many of our entity representations are shared with others, in the sense that they too have an entity representation for the entity that is represented by the one we have. What it means for an entity representation of mine and an entity representation of yours to 'share their reference' is something that needs proper spelling out, and the spelling out is crucial for what I want to say in the last part of this chapter.

At the end of Section 2 I observed that the internal coreference between the two beliefs of the mental state description given in (1) wasn't a good illustration of the problems connected with internal coreference because the shared dref x represented an existing entity (the actual disease in question). We can now make this observation formally precise. To this end we must not only assume that cancer of the pancreas is an existing disease (which it is), but also that our agent John knows that it is and that this knowledge takes the form of an ER which represents the disease and that has one or more anchors testifying to the causal link or links that connect it with the disease. On these assumptions, John's mental state contains an ER for the disease, and a more accurate description of it than (1) would be the one in (2).

$$(2) \left\{ \begin{array}{l} \left\langle [ENT, x], \begin{array}{|c|} \hline \\ \hline \text{disease}(x) \\ \hline \text{Named}(x, \textit{cancer o.t. panc.}) \\ \hline \end{array}, \mathcal{K} \right\rangle \\ \\ \left\langle BEL, \begin{array}{|c|} \hline s \\ \hline n \subseteq s \quad \text{canc.o.t.pancr.}'(x) \\ s: \text{have}'(i, x) \\ \hline \end{array} \right\rangle \\ \\ \left\langle BEL, \begin{array}{|c|} \hline s' \quad d \\ \hline n \subseteq s' \quad \text{drug}'(d) \\ s': \text{cure-for}'(d, x) \\ \hline \end{array} \right\rangle \\ \\ \left\langle DES, \begin{array}{|c|} \hline e \\ \hline n \prec e \\ e: \text{obtain}'(i, d) \\ \hline \end{array} \right\rangle \end{array} \right\}$$

The mental state representation in (2) captures the ‘aboutness’ of the beliefs in question by means of an ER, i.e. a special type of mental state component identified with the mode indicator ENT, introducing a dref x used in both belief components.

The general definition of ERs we will be using is given in (3):⁴

- (3) An *Entity Representation* is a triple $\langle [ENT, x], K_{descr}, \mathcal{K}_{anch} \rangle$, where
- (i) ‘ENT’ is a ‘Mode Indicator’, which indicates that the constituent is an entity representation, as distinct from the various types of propositional attitudes, beliefs, desires etc, that also populate the

⁴Since ERs were first introduced, their form has been subject to some variation (see Kamp (1990, 2003)). The form defined in (3) that is adopted here has proved adequate for a number of different purposes and I do not expect that further modifications will be needed soon, if at all. To those familiar with the notion of ‘file cards’ or ‘mental files’ (e.g. Heim 1982,1988; Recanati 2012; Maier 2016) the similarities between ERs and file cards will be obvious. But there are also significant differences between ERs and ‘file cards’. The term ‘Entity Representation’ was chosen to prevent unintended assumptions about ERs on the part of those who are familiar with one or more file card theories.

mental states of MSDRT;

- (ii) x is a discourse referent (the *distinguished dref* of the ER);
- (iii) K_{desc} is a DRS (which contains certain types of descriptive information about the represented entity);
- (iv) $\mathcal{K}_{anch}$ is a (possibly empty) set of *internal anchors* (internal anchors are DRSs from some special purpose DRS language, an *anchor language*).

The second component of the ER provides descriptive information about the ER's referent in the form of a DRS. Anchored ERs are causally tied to their referents via the relation or relations witnessed in their anchor set $\mathcal{K}_{anch}$. When the anchor set is empty, however, it is the second component that serves to determine the referent of the ER, as the entity that uniquely fits the description provided by this component. (For more on this see Section 3.3.) For anchored ERs there appears to be a certain flexibility as to what descriptive information goes into their second component and what should take the form of separate beliefs.

The third component of an ER contains the anchors, which act as witnesses, at the level of internal representation, of how the ER is causally linked to its referent. Anchors carry information about the nature of the causal links between their ERs and the entities they represent. Perceptions are the prototypical kinds of causal linking. A perception can either give rise to the creation of a new ER in the mind of the perceiving agent, with an anchor that testifies to its perception-based origin and links it to the entity perceived; or, alternatively, the perception may evoke an ER that already exists in the perceiver's mind, in which case a perceptual anchor is added to the anchor set of this ER as a witness to the current perception; these are the cases where the perceiver recognizes what he is currently perceiving as something he already knows.⁵

The anchor set of the ER in (2) has been left unspecified, in the form of the schematic $\mathcal{K}$. The presumption here is that the ER does have one or more anchors that witness its causal connection with the disease, but without saying anything about the form of these anchors. But let us ask: What could such an anchor be like? Here is a partial answer: First, the anchor could be

⁵One of MSDRT's assumptions is that recognizing an entity as something that you already know always takes this form – of identifying the entity you have encountered with the referent of an ER that you already have and adding a new anchor to the anchor set of this ER as a witness of your present encounter with the entity.

a perceptual one, witnessing a perception by the agent of the represented entity. In the case before us a perceptual anchor is perhaps not a particularly prominent possibility. What could it mean to ‘perceive a disease like cancer of the pancreas’. Arguably such a thing is possible for people from the medical profession, who may come into actual visual contact with the disease when they see its manifestations in particular patients, e.g. by looking at their X-rays. But that doesn’t apply to most of us, who do not deal with diseases from a professional stance (though of course those who suffer from the disease may also stand in some perceptual relation to it, by feeling the detrimental effects it has on their health). But for a lay person it is more common to know about the disease because they have heard about it, or because someone has told them about it or because we have read about it somewhere. (And presumably that is how doctors first become acquainted with the disease too. They learn about it from lectures or books in the course of their medical education.)

One of the central assumptions of MSDRT is that such encounters – with mentions of a referent that occur in texts or that are made by people to whom we are listening – also establish relations to the mentioned referent that enable us to entertain thoughts about it and to express these thoughts in language. In other words, they enable us to create an ER for the thing encountered, with an anchor that witnesses the way in which this ER is causally linked to its referent via the mention whose interpretation prompts the ER’s creation or reuse. Anchors of this second kind, which witness the agent’s commitment to treat her ER as representation of whatever it was that the referring expression she has just interpreted referred to, are known as *vicarious anchors*. When my doctor tells me that I have cancer of the pancreas and I do not doubt that what he tells me is true, then from that point onwards I will have an ER for this disease with a vicarious anchor that reflects my getting information about the disease from him. Perhaps this is the first time I hear about the disease, of which until then I didn’t know that it existed. In that case I will form a novel ER to represent the disease, with a single vicarious anchor that testifies to what prompted its formation. But it is also possible of course that I already knew about the disease and already had an ER to represent it. In that case my interpretation of what the doctor tells me will make use of this ER, and augment its anchor set with the vicarious anchor spoken of above.

Perceptual anchors and vicarious anchors are the only anchor types that will matter in this chapter. Vicarious anchors will play a particularly prominent part and it will therefore be useful to say a little more about what such

anchors look like. To this end, suppose that our agent John of (2) first learned about cancer of the pancreas from what he has been told by his doctor and that his ER for this disease is as in (4).

$$(4) \quad \left\langle [ENT, x], \begin{array}{|c|} \hline \\ \hline \text{disease}(x) \\ \hline \text{Named}(x, \text{ca'r o.t.panc.}) \\ \hline \end{array}, \left\{ \begin{array}{|c|} \hline e \\ \hline e \prec n \\ \hline e: \text{ref}(z, \text{ca'r o.t.panc.}, x) \\ \hline \end{array} \right\} \right\rangle$$

The form adopted here for the vicarious anchor makes use of the predicate ‘ref’. ‘ref’ is an event predicate with four argument slots. The 1st slot is for events, the 2nd slot is for the information source z from which the agent adopts the referent, the 3rd slot is for the expression that was used by or in the source to refer to the referent (in the case of (4) this is assumed to have been the phrase *cancer of the pancreas*), and the 4th slot is for the referent itself, and is always filled by the distinguished dref of the ER to which the anchor belongs.

Note that the dref z for the information source that fills the 2nd argument slot of the predicate ‘ref’ in (4) isn’t bound in its DRS. In the case we are looking at the binding is provided by another ER, the one that the agent has for the information source – in this case the doctor from whom he has learned about his plight. (5) is a revised version of the mental state description in (1), in which it has been extended with both an ER for the disease and an ER for the doctor from whom John, the possessor of the state, learned about the existence of the disease.

$$(5) \left\{ \begin{array}{l} \left\langle [ENT, z], \begin{array}{|c|} \hline \\ \hline \text{doctor-of}(z, i) \\ \hline \end{array}, \mathcal{K}_z \right\rangle \\ \left\langle [ENT, x], \begin{array}{|c|} \hline \\ \hline \text{disease}(x) \\ \text{Named}(x, \text{cancer o.t. panc.}) \\ \hline \end{array}, \left\{ \begin{array}{|c|} \hline e \\ \hline e \prec n \\ e: \text{ref}(z, \text{ca'r o.t. panc.}, x) \\ \hline \end{array} \right\} \right\rangle \\ \left\langle BEL, \begin{array}{|c|} \hline s \quad x \\ \hline n \subseteq s \quad \text{canc.o.t.pancr.'}(x) \\ s: \text{have}'(i, x) \\ \hline \end{array} \right\rangle \left\langle BEL, \begin{array}{|c|} \hline s' \quad d \\ \hline n \subseteq s' \quad \text{drug}'(d) \\ s': \text{cure-for}'(d, x) \\ \hline \end{array} \right\rangle \\ \left\langle DES, \begin{array}{|c|} \hline e \\ \hline n \prec e \\ e: \text{obtain}'(i, d) \\ \hline \end{array} \right\rangle \end{array} \right\}$$

None of the ERs shown in (1), (2) and (5) guarantee that there is an entity they represent. An internal anchor can be the result of a mistake. For example, the visual perception that leads to the formation of an ER for what the agent takes himself to be seeing may be a visual illusion: there is no object there to be perceived. In that case the perceptual anchor of the ER is a false witness and the ER doesn't represent anything. Likewise, what a listener or reader takes to be a properly referring use of an expression need not be that – the speaker or author may have been under an illusion that she was making a proper reference, or she may have pulled the recipient's leg.

3.2 Types of Anchors and Absence of Anchors

There are two types of anchors we have considered, perceptual anchors and vicarious anchors. These differ from each other in that a perceptual anchor witnesses a direct encounter with the referent of its ER, through a causal perception of it by the possessor of the mental state of which the ER is a constituent; a vicarious anchor, on the other hand, is witness to an encounter with a *reference to* the referent, made by an information source that can be assumed to stand in some direct or indirect causal relation to the referent.

References made by other sources can take many different forms. In this chapter we focus on the cases where the reference is made with the help of a referring expression, and more particularly those where that expression is a proper name.

It could be held that vicarious anchors are perceptual anchors of a kind; they witness events in which the possessor B of the ER to which the anchor belongs observes, visually and/or acoustically, some source A that B assumes stands in some suitable causal relation to the referent $\mathbf{d}$. On the strength of this observation B can then take himself to also stand in a suitable causal relation to $\mathbf{d}$, a causal relation that is the composition of (i) the presumed causal relation between A and $\mathbf{d}$ and (ii) the causal relation between A and B consisting in A 's reference to $\mathbf{d}$ by using some referential expression α and B perceiving A 's use of α as a case of referring (to some entity). According to MSDRT this compound causal relation between B and $\mathbf{d}$ qualifies as a 'direct reference relation', which enables B to entertain thoughts and make utterances that are about $\mathbf{d}$; in MSDRT terms this means that B is in a position to form an ER that represents $\mathbf{d}$.

As noted, our definition allows for ERs with empty anchor sets. It is one of the stipulations of MSDRT that such ERs refer by virtue of the descriptive information in their second component: the ER refers if there is a unique value for its distinguished dref x for which its descriptive component is true; otherwise the ER is improper and fails to refer. Or, more formally and precisely, given an intensional model M with actual world w_0 , the ER is *proper with respect to M* iff there is a unique $\mathbf{d}$ in the Universe of M_{w_0} that satisfies the descriptive component of ER; and if not, then, relative to M , the ER is defective and without a referent. (These reference conditions for anchor-free ERs will remind the reader of the traditional Fregean account of the reference conditions of definite descriptions. Indeed, according to MSDRT the interpretation of a definite description may involve an anchor-free ER, either a new one or one that was part of the interpreter's mental state already. We will return to this point presently.)

3.3 Singular Content

This is a good place for a general observation about the model-theoretic semantics of the DRS languages of MSDRT that could have been inserted earlier, but that has to be made at some point in this chapter. It is a very common feature of the DRSs of MSDRT that the propositions they express

are *singular propositions*. This is so in particular for the DRSs that specify the contents of attitude constituents in ADF descriptions of mental states, but it is true more generally, partly because of the fact that many DRSs contain ADF descriptions as constituents.

First an informal reminder of what is meant by ‘singularity’ and a formal specification of how the term will be used in this chapter. A proposition is singular with respect to some entity $\mathbf{d}$ iff it attributes some property P to $\mathbf{d}$: the proposition is true in any possible world w iff $\mathbf{d}$ has P in w . Here follows a formal definition of singularity, within the setting of model-theoretic semantics, and more specifically, that of the model theory of the DRS languages of MSDRT. The definition requires a few words about the model theory for such DRS languages. It is an *intensional* model theory, in which the models are possible world models. A possible world model M is based on a set W_M of possible worlds and provides for each world w in W something like an extensional model, which gives the denotations of terms and the extensions of predicates in that world. One of the worlds in W_M , which we will refer to as w_0 , plays the role of the actual world. A *proposition p relative to M* is a subset of W_M (intuitively: the set of those worlds of M in which the proposition is true). A proposition p relative to M is *singular with respect to* the individuals $\mathbf{d}_1, \dots, \mathbf{d}_n$ iff there is an n -place relation R such that p is the set of those worlds w in W_M that are such that R is satisfied by $\langle \mathbf{d}_1, \dots, \mathbf{d}_n \rangle$ in (the extensional model associated with) w . Intuitively this is the case when p can be thought of as being *about* the entities $\mathbf{d}_1, \dots, \mathbf{d}_n$: it says about them that they stand in some relation R to each other.

Many of the content specifications of propositional attitude components of AFD descriptions determine singular propositions. This is so whenever a content specifying DRS K contains the distinguished dref x of a properly referring ER ER in the given AFD description AD . Intuitively the reason should be clear. K should evaluate as true if and only if the entity $\mathbf{d}$ represented by ER satisfies what K predicates of x . In other words, the proposition expressed by K is a singular proposition about $\mathbf{d}$. (More generally, if $x_1, \dots, x_n$ are the distinguished drefs occurring in K , then K expresses a proposition that is singular with respect to $\mathbf{d}_1, \dots, \mathbf{d}_n$, the entities represented by the ERs $ER_1, \dots, ER_n$ that have $x_1, \dots, x_n$ as their distinguished drefs.)⁶

So far in this discussion of propositional singularity I have tacitly assumed

⁶This informal argument that many of our attitudes have singular contents is confirmed by the model theory of MSDRT. But the details cannot be adequately explained within the confines of this chapter. More about these formal details can be found in the references

that ERs are properly anchored. For ERs with non-empty anchor sets this will be the case when the anchors in the set link the ER causally to a single entity. But what about ERs with empty anchor sets? According to definition (3) such ERs refer to the unique satisfiers of the descriptive content in their second component. That may suggest that such ERs function essentially like definite descriptions, in either a Fregean or a Russellian spirit. But that is not how such ERs are assumed to function in MSDRT. The role of an ER whose referent is determined by its second component is more like that of a proper name that has been introduced by means of a description, as discussed in Kripke (1980). For Kripke the name *Neptune*, which was originally introduced as name for the planet that caused certain perturbations in the orbit of Uranus, functions like other proper names in that in its normal uses it refers directly, introducing its referent directly into the propositional content of sentences containing it. An ER with empty anchor set functions analogously, directly contributing its referent – i.e. the unique satisfier of the descriptive content that is specified by its second component – to the propositional content representations elsewhere in the mental state that contain occurrences of the ER’s distinguished dref (as in the content-DRSs of the beliefs in (2) and (5)). Thus when the content specification K of an attitude in a given AD contains the distinguished dref x of an ER with empty anchor set, and

mentioned in footnote 1. In a nutshell the story is this. (But there is no need to fully absorb this to understand what follows in the chapter.) The extensional models M_w of MSDRT’s intensional models M may contain agents and specifications of their mental states. The specification in M_w of the mental state of agent $\mathbf{A}$ will as a rule include causal relations between $\mathbf{A}$ and entities belonging to the universe of the model M_w which enable $\mathbf{A}$ to entertain thoughts about these entities. Furthermore, the contents of many of the attitudes that are part of $\mathbf{A}$ ’s mental state in M_w are specified as thoughts about entities to which $\mathbf{A}$ is causally related, that is as propositions that are singular with respect to those entities. According to the model theory an ADF description AD *correctly describes* $\mathbf{A}$ ’s mental state in M_w iff the following condition is satisfied: There exist

- (i) an assignment r from ERs in AD to entities in the set $\{\mathbf{d}_1, \dots, \mathbf{d}_n\}$ of entities in M_w to which $\mathbf{A}$ is causally related as specified by the model M_w ;
- (ii) an embedding g that maps the propositional attitude components of AD to propositional attitudes of $\mathbf{A}$ ’s mental state in M_w which preserves attitudinal mode – i.e. beliefs from AD are mapped to beliefs of $\mathbf{A}$ ’s mental state, desires are mapped to desires and so forth – such that when K is the DRS that specifies the content of a component $\langle \text{MOD}, K \rangle$ of AD , the proposition that K expresses relative to r is entailed by the propositional content of $g(\langle \text{MOD}, K \rangle)$. Here the *proposition that K expresses relative to r* is determined in the usual way for DRSs in intensional models: the proposition is true in any extensional model M'_w of M iff there is a verifying embedding f of K in M'_w which maps each distinguished dref x of an ER ER in AD to the entity $r(ER)$.

the ER's descriptive content has a unique satisfier $\mathbf{d}$, then the proposition expressed by K is a singular proposition about $\mathbf{d}$, just as it would be if x had been the distinguished dref of a properly representing ER with non-empty anchor set.

For an illustration, suppose that agent A uses the sentence 'The first man to be born in the 22nd century will be born in Hong Kong' to express a belief $\langle \text{BEL}, K \rangle$, where K is of the form ' x will be born in Hongkong'⁷ and x is the distinguished dref of an ER with empty anchor set and a descriptive component to the effect that x will be the first person to be born in the 22nd century. Then the propositional content of this belief of A 's will be a proposition that is singular with respect to the first person to be born in the 22nd century – provided there will be such a person; and if there won't be, then the content of A 's belief is undefined. And if A 's utterance is to be accepted as an expression of this belief, then it too should get an analysis in which it expresses this singular proposition and in which the description 'the first person to be born in the 22nd century' is used 'referentially'. (Perhaps it is impossible to entertain singular propositions of this sort, and to use such future-related descriptions referentially, but that is another matter.)

3.4 Causal chains and ER networks

According to MSDRT ERs play an important role in verbal communication. When the recipient of a spoken or written utterance takes a noun phrase α occurring in it as *referential*, i.e. as directly referring to some particular entity, he will look for an ER in his entity library that, as far as he can tell, represents the entity referred to. If he can find such an ER, he will use it (via its distinguished dref) to represent his interpretation of the utterance and he will add a vicarious anchor witnessing his use of this ER for his interpretation of α to its anchor set. And when he cannot find a suitable ER, then he will accommodate by creating a new one, with a vicarious anchor as witness (recall the discussion of vicarious anchors in Section 3 and their form given in (4)).

But ERs play an equally important part in utterance production. When the mental state of an agent A at time t contains an ER ER then that puts A in a position to use a referential noun phrase α to refer to the entity

⁷Presumably A will also have an ER for Hongkong, in which case her belief will be doubly singular, with respect to both the first person born in the 22nd century and to Honkong. For more about proper names see 3.5 and much of what follows after that.

represented by ER . According to MSDRT, using an expression α to refer to the entity represented by an ER ER in one's current mental state is a primitive concept, which cannot be reduced to simpler notions. Each such referring use is the manifestation of this irreducible 4-place relation between agent, ER, referent $\mathbf{d}$ and the expression α that is being used.⁸

Suppose that A utters α to refer to the entity represented by an ER ER_A . And suppose that her interlocutor B interprets this utterance by adding a vicarious anchor that witnesses his interpretation of α to his representation ER_B for α 's referent. So far I have talked about this vicarious anchor as establishing a link between ER_B and the referent of ER_A . But by the same token, the anchor can also be seen as establishing a link between the two ERs ER_A and ER_B . Such connections, which are formed or confirmed each time one speaker refers to some particular entity and someone in her audience takes her to be doing that, are the links of networks that connect the ERs belonging to different speakers, which grow larger and more complex as time goes on. I will refer to these links between ERs of different speakers as *coreference links*.

3.5 Named ERs and Referring by Name

Among the different noun phrase types that can be used to establish or sustain coreference links proper names play a special part, and one of unrivaled prominence. To get a sense of this, here is a simple example to start with. Suppose that Susan says to Bill:

(6) Mary is in Paris.

As argued in ?, the kind of use of proper names illustrated by *Mary* and *Paris* in (6) – that where English proper names form DPs all on their own – presupposes that Susan has ‘Named’ ERs for the entities to which she refers by using these names. A ‘Named’ ER is one whose descriptive component contains a ‘Naming Condition’ – a DRS Condition of the form ‘Named(x , N)’, where N is a name and x is the distinguished dref of the ER to which the Condition belongs. This Condition says of the referent of its ER that it

⁸The predicate ‘ref’ occurring in (4) is closely related to this 4-place relation. When the vicarious anchor in (4) is the result of interpreting a phrase α uttered by an interlocutor A , then the interpreter who adds (4) to the relevant ER ER will assume that this relation holds between A , the ER that A has used in her utterance of α , the referent of his own vicariously anchored ER and the referring expression α .

goes by the name N . I will refer to an ER with the Condition ‘Named(x, N)’ also as ‘ N -labeled’ and also to ‘ N -labeled’ ERs more simply as ‘labeled ERs’ or ‘named ERs’. Note that an ER can be N -labeled for more than one N : Entities can have more than one name and the possessors of ERs representing such entities may know that.

It will be assumed from here on that a referential use of a name N is always based on an N -labeled ER that the user has for the referent to which she is using N to refer. (Indeed, how else could a user make a referential use of a name? How could she use the name if she didn’t know it was a name for what she wants to refer to? In our set-up the only form this knowledge can take is that of a named ER for the referent.) In short, then, (6) presupposes that Susan has a *Paris*-labeled ER for Paris and a *Mary*-labeled ER for the Mary (6) is speaking of.

We now turn to the addressee Bill. Suppose that Paris is familiar to Bill but that he has no idea who the Mary is that Susan is talking about. That is, Bill too has a *Paris*-labeled ER for Paris, but to the best of his knowledge he doesn’t have a *Mary*-labeled ER for the relevant Mary. Under those conditions Bill may be expected to use his ER for Paris in his interpretation of Susan’s use of *Paris* and react to Susan’s utterance of *Mary* by setting up an ER for the Mary that he assumes Susan is referring to, but about whom, he so far doesn’t know anything else. And in both cases Bill will introduce a vicarious anchor into the anchor sets of these respective ERs as witnesses of his interpretations of the occurrences of the names *Paris* and *Mary* in (6). These vicarious anchors establish the coreference link spoken of above between Susan’s and Bill’s ERs for Susan’s use of *Paris*, and that between the Susan’s pre-existing and Bill’s new ER for Susan’s use of *Mary*.⁹

(7) displays the relevant part of the mental state of Bill after interpreting (6).¹⁰

⁹It is of course possible that one of the ERs that Bill already had does in fact represent the Mary that Susan is referring to. In that case Bill will now have two ERs that both represent this Mary, but without knowing that they do represent the same individual (or perhaps he has even more than two, but let that pass). It is even possible that both of these are *Mary*-labeled. But in that case Bill will think that they represent different Marys.

¹⁰A reviewer asks for clarification of the use of subscripted symbols like ‘ s_B ’ in (6). Brief answer: The ERs in (6) are representations in the mental state of Bill for the different individuals Sue, Mary and Paris. The symbols ‘ s_B ’ etc for the distinguished drefs of these ERs have been chosen for mnemonic reasons: the s in ‘ s_B ’ is an informal pointer to Susan, and the subscript ‘ $_B$ ’ is to put us in mind that ‘ s_B ’ is the distinguished dref of an ER that is part of the mental state of agent Bill. Formally, what matters is only that distinct dref

$$(7) \left\{ \begin{array}{l} \left\langle [ENT, s_B], \begin{array}{|c|} \hline \\ \hline \text{Named}(s_B, Susan) \\ \hline \end{array}, \mathcal{K}_{s,B} \right\rangle \\ \left\langle [ENT, m_B], \begin{array}{|c|} \hline \\ \hline \text{Named}(m_B, Mary) \\ \hline \end{array}, \left\{ \begin{array}{|c|} \hline e_m \\ \hline e_m \prec n \\ e_m: \text{ref}(s_B, Mary, m_B) \\ \hline \end{array} \right\} \right\rangle \\ \left\langle [ENT, p_B], \begin{array}{|c|} \hline \\ \hline \text{Named}(p_B, Paris) \\ \hline \end{array}, \mathcal{K}_{p,B} \cup \left\{ \begin{array}{|c|} \hline e_p \\ \hline e \prec n \\ e_p: \text{ref}(s_B, Paris, p_B) \\ \hline \end{array} \right\} \right\rangle \\ \left\langle BEL, \begin{array}{|c|} \hline s'_B \\ \hline n \subseteq s'_B \\ s'_B: \text{in}'(m_B, p_B) \\ \hline \end{array} \right\rangle \end{array} \right\}$$

The link between the ERs for Paris in Susan's and Bill's mental states that is established by the displayed vicarious anchor in (7) may be part of a longer chain. For instance, it is possible that neither Susan nor Bill has ever been to Paris. Their knowledge is only indirect, they heard about it from others; and so the only anchors of their ERs for Paris are vicarious ones¹¹. This means that their *Paris*-labeled ERs refer only on the strength of being linked to the *Paris*-labeled ERs of others, and these possibly in their turn to yet others. All these ERs representing Paris form a network within the speech community to which their possessors belong. It is a network that grounds the *common knowledge* of Paris that is shared between the possessors of these ERs.

As I have described the exchange (6) between Susan and Bill, their respective *Paris*-labeled ERs are already linked, either directly, through previous occasions when the two of them talked about Paris, or indirectly, because both talked at some time or other about Paris to some other person, or to other persons who had been communicating about Paris. In such cases the new link between the two ERs that is documented by the vicarious anchor

symbols are the graphic realizations of distinct drefs, for instance that ' s_B ' is a different dref from ' m_B '.

¹¹In all likelihood a whole bunch of them, since most of us are exposed to *Paris* as the capital of France over and over again, even if we ourselves have never been there.

which the interpreter adds to his ER for the referent does no more than reinforce an already existing connection between them. The story just told about Bill’s interpretation of *Mary* is different in this regard. Bill creates a new *Mary*-labeled ER for the Mary that Susan has referred to. The vicarious anchor of this ER links it to the ER that Susan has used in her utterance of *Mary*. But at this point this is the only way in which his ER is linked. On the other hand, Susan’s ER is presumably linked to Mary-representing ERs of other users and thus already part of a larger network of such ERs.¹² The newly established link between her ER and Bill’s integrates his ER into this network.

Our story about the production and interpretation of (6), with its ramifications for networks of Named ERs within the speech community, is in outline MSDRT’s reconstruction of the causal chain account of names of Kripke and Chastain (??): Suppose that B ’s N -labeled ER ER_B for an entity $\mathbf{d}$ was created in response to a reference that A made by uttering the name N , based on her N -labeled ER ER_A for $\mathbf{d}$. If ER_A properly represents $\mathbf{d}$, then so will ER_B , by virtue of the vicarious anchor that links it to ER_A . For ER_A we can distinguish two possibilities: either (i) A was present at the event when N was bestowed on $\mathbf{d}$ as a name (by an act of ‘baptism’ as Kripke puts it) or (ii) she wasn’t. In case (i) B ’s use of N has been traced back to the creation of the name through a single coreference link. In case (ii) this is not so. Here ER_A will have one or more vicarious anchors that connect it with N -labeled ERs of yet other members of the community. With regard to those ERs the same question arises as to how they represent $\mathbf{d}$, and so on. The present version of the causal chain theory is that at least one of the backward chains that can be formed in this way reaches a participant of the baptismal introduction of N after some finite number of steps.

As a rule there will be many chains that lead from a given N -labeled ER ER_B back to an ER of someone present at the baptism. This makes no difference to the legitimacy of B using N – one chain would suffice. But the number of different backward chains is significant nonetheless. This number determines what we might loosely call the ‘density’ of the network, some measure of how widely and regularly a name is used within the community. Density, I will argue in Section 4, is especially important in connection with names in fiction.

¹²And it is also possible of course that she knows the Mary she is referring to personally, so that her ER will have one or more perceptual anchors.

The density of the network of N -labeled ERs for a given referent can be seen as an indicator of the entrenchment of a name N . But there is also a flip-side to density. Anchored ERs are vulnerable to certain forms of corruption. One form was already mentioned: An ER may be based on a fake perceptual experience, where there is nothing that is actually perceived. In such cases the ER doesn't represent anything and no propositional representation that makes use of it (by containing one or more occurrences of its distinguished dref) determines a well-defined proposition¹³. But multiply anchored ERs are also threatened by another source of corruption: they can become *incoherent*, when two or more anchors in their anchor set connect them with different referents. And that, I submit, is just as bad. Being linked to more than one referent isn't much better than being linked to none at all.

The larger the anchor set of an ER, the greater the threat of incoherence. Just one faulty recognition in the course of its history – one occasion when the agent thinks it is Susan he sees going by on the other side of the street, but it isn't her – and damage has been done. Moreover, when the ER is part of an extended network, as so many of them are and named ERs practically always are, then the damage may not be limited to this one ER, but may spread through the network, causing global as well as local disruption. And the denser the network, the bigger the risks of spreading.

Yet we shouldn't overdramatize. Real as the threats to coherence may be, the actual danger they carry may nevertheless keep within limits. A single deviant anchor need not be the end of an ER's practical usefulness. In fact, ERs seem to be remarkably robust to the nefarious effects of incoherence, especially when their anchor sets are large. As often as not an agent will soon forget about the faulty uses he has made of a multiply anchored ER – the woman on the other side of the street may disappear beyond his cognitive horizon almost as soon as she made her brief appearance within it and no significant trace remains of the error he made when he took her for Susan. Because of this robustness of ERs I will set the problem of incoherence, and in fact of reference failure of ERs more generally, aside in what follows. That is an idealization of course, but it doesn't seriously affect what I want to say about networks of coreference links, and especially not about coreference networks of ERs labeled by the same name.

¹³Unless the ER has other anchors that make up for this failure. This possibility is connected with the robustness of impaired ERs mentioned below. As explained there, we set such cases aside in this chapter.

3.6 Learning Names from Introductions and from Texts

From what has been said so far it might look as if the main mechanism by which a name spreads throughout a speech community is accommodation. When the recipient of an utterance containing a name *N* has no ER for the reference of that name, then he has to accommodate by introducing a new ER for its referent, which gets linked by its vicarious anchor to the ER that was used by the speaker. But the impression I may have created is somewhat misleading. Not all uses we make of names fit this description of their production and interpretation. It only fits what in Kamp (2015) is called the ‘standard’ use of names. But names have other uses besides their standard use. In particular, they can be used in introductions, as when Susan says to Bill: “Meet my friend Mary”, while clearly referring to the woman standing next to her. In this case Bill isn’t expected to have had an ER for the referent of the name before Susan made her utterance. It is in keeping with the purport of her use of *Mary* for him to end up with a *Mary*-labeled ER for the referent then and there. This may either take the form of adding *Mary* as label to an ER that he already has of the friend (e.g. by having noted her before Susan said anything), or by introducing a completely new ER into his ER library, in case it is only Susan’s utterance that brings her friend to his attention.

Another example is when a teacher says to his high school students “Today we are going to talk about a Greek philosopher who has had a greater impact on science than anyone else in history. His name was Aristotle.” Here the teacher’s first sentence enables her students to set up an ER for the Greek philosopher it mentions. In response to the second sentence they will then add to this ER the label *Aristotle*.

Such introductory uses of names clearly contribute to their spread within the community as well. That’s what introductory uses are for. But perhaps the most important factor in the spreading of names, especially if names for famous entities such as *Aristotle*, is the occurrence of names in texts.¹⁴ In principle, the standard use of names in texts doesn’t differ from their standard use in face to face communication. Standard uses of names in texts convey, just as the standard uses of names in spoken utterances, the presumption that they are familiar to the reader – in our terms, that the reader has a labeled ER for their referent that he can use for their interpretation. But

¹⁴For now I am thinking just of non-fictional names occurring in public non-fictional texts. Fictional names – names of fictional entities – are the topic of Section 4.

in practice the way names in texts are handled tends to be quite different. Authors cannot attune the way they express themselves as closely to the needs of individual readers as speakers can who have some information about the knowledge and ignorance of those they are addressing. As a rule an author has only limited knowledge of who will read her texts. When the author needs to use a name N , then using it in the standard way would be appropriate to readers familiar with N and its referent, but perhaps less than considerate to those for whom the name is new. And conversely, making an introductory use of the name might be more considerate to readers of the second kind but come across as condescending to readers of the first kind. There seems to be a tendency for authors to opt for the standard use when they are in doubt about their readership (or should be). And this we acknowledge as readers. When we encounter a standardly used name in a text that is not familiar to us, our natural attitude is to accommodate, and to see it as part of our task of understanding the text to find out more about the referent – in the terms of MSDRT: We are ready to adopt a new labeled ER for the referent, and make that the crystallization point for what more we will find out about the referent from then on. In this way new ERs become the hubs of our learning.

But if the interpretation of names in texts plays such an important part in the establishment of coreference networks, what are the ERs like that readers accommodate for the unfamiliar names that they encounter in texts? In particular, what in these ERs makes them part of the networks for the names they interpret? Do these ERs also have vicarious anchors, like the ERs involved in the interpretation of names occurring in spoken utterances?

There are two aspects to this question. First, there is the question how we can suppose a reader to record his interpretation of a name encountered in a text, and that is either familiar or unfamiliar to him. And second, there is the question how this record integrates his ER into the existing network for the interpreted name (in case the name is unfamiliar), or reinforces the position it already occupies in this network (in the case of familiar names). I'll take these two questions one by one.

The story I have told about the accommodation of unfamiliar names in utterances has it that the accommodated ERs get a vicarious anchor that records the reason for their construction. And such a vicarious anchor is also added to an ER that the interpreter already had in those cases where accommodation is not needed. As shown in (4), such a vicarious anchor specifies the producer of the utterance as the one responsible for the ER. Something like this might also work for ERs that are used or introduced

in response to unfamiliar names encountered in texts, especially when the reader knows about the authorship of the text. But even in such cases the situation of the reader is different from that of someone who interprets a spoken utterance at the time when it is made. We read texts after, and quite often long after, they have been produced. And often we do not even know who the author was; for all we know it may even have been some kind of anonymous collective. In such cases all that is available to us when we encounter a name in the text is that this name is occurring in the text at the particular point where we encounter it. I will assume that it is just this information that we record as part of our interpretation of the name, that this record takes the form of a kind of vicarious anchor in the anchor set of the ER we make use of in our interpretation and that this anchor links the ER to this given occurrence of the name.

The way in which this assumption can be implemented is shown in (8) below. But before we get to that we should have a look at the second aspect of our question: What is the effect of such anchors on coreference linking and network expansion? What are the other ERs to which these anchors link the ERs of which they are constituents? At an intuitive level it would seem quite clear what the answer should be: The producers of texts must have used the names that made it into their texts to refer to what they took those names to refer to when they wrote. They had labeled ERs for the referents of those names, and it is only because of this that it makes sense to take the names in the text as referring. That is what any normal reader implicitly takes for granted.

This answer implies that the vicarious anchor of the ER used by the reader of a written token of a name N in a text T links it to the ER that the author or authors used when they entered this token into their text. But the formal interpretation of this conclusion raises some further questions, which it is important to turn to briefly. The issues we will have to touch upon to in this connection are quite complex and I won't be able to do justice to them. One simplification I will make is that texts have unique authors. That of course isn't true in general, but the simplification will be harmless: it should not be too hard to see how the things I will say can be adapted to the case of multiple authorship.

The idea underlying the formalization of the coreference linking of ERs used in the interpretation of names in texts is that texts are among the players that make up a language community, with its language and the wider culture within which community and language are embedded and that the

language both reflects and partly defines. On the one hand texts may look like auxiliary players, as go-betweens that connect authors and readers. But from a historical perspective their role is more than that. Texts are born for longevity, which is free of the restrictions that biology imposes on the human members of the language community. This enables them to act as bridges to distant pasts that no single human existence can bridge. It is because of their longevity that many texts are often our primary connection with a distant past populated with creatures and things long gone and that would have been expunged from memory it hadn't been for those texts, and the names they contain for what otherwise would have been lost for good.

So much for a bit of background. The formal treatment of the coreference links established by ERs used in the interpretations of names in texts that I propose in this light is this as follows. I assume that with each token N_i of a name N occurring in a text T is associated a *pseudo-ER*, an ordered pair consisting of N_i and the referent of the ER that the author relied on when she wrote N_i at the time, as part of writing T . This pseudo-ER will thus be linked on the one hand to the ER involved in the production of N_i , by taking over the referent of that ER. In what follows I will assume that these links between the ERs used by the author and the pseudo-ERs associated with the names she used, are part of the coreference networks for those names. On the other hand we will assume that it is the pseudo-ER that are the direct link mates for the ERs that are used by readers to interpret the names of T . It is through these two kinds of links that pseudo-ERs act as intermediaries between the name-labelled ERs of authors and the name-related ERs of their readers.

To sum up, each referential use of a name token in a text is associated with a pseudo-ER, which consists of a unique identifier of the token and the referent of that token. It is this pseudo-ER that is referred to in the vicarious anchors to the ERs that readers of the token use in their interpretation of it. With this we return to the form of these vicarious anchors.

The point of the representation in (8) is the anchor of the third ER. The other two ERs, for the name token ν whose interpretation results in the third ER and the text τ in which ν occurs, are shown because they are mentioned in this anchor.

$$(8) \quad \left\{ \begin{array}{l} \left\langle [ENT, \tau], \begin{array}{|c|} \hline \text{text}(\tau) \\ \hline \end{array}, \mathcal{K}_\tau \right\rangle \\ \left\langle [ENT, \nu], \begin{array}{|c|} \hline \text{name-token}(\nu) \\ \hline \text{In}(\nu, \tau) \end{array}, \left\{ \begin{array}{|c|} \hline s \\ \hline n \subseteq s \\ s: \text{see}(i, \nu) \end{array} \right\} \right\rangle \\ \left\langle [ENT, x], K_\nu, \left\{ \begin{array}{|c|} \hline s \\ \hline n \subseteq s \\ s: \text{ref}(\tau, \nu, x) \end{array} \right\} \right\rangle \end{array} \right\}$$

Some assumptions relating to the anchor of the third ER in (8):

(i) The predicate ‘ref’ is a variant of the predicate ‘ref’ of (4) which differs only in the sorts of its arguments. The predication ‘ref(τ, ν, x)’ says that the text τ uses the name token ν to refer to the entity x that is represented by the ER of which the predication is part. A further difference is that the ‘ref’-predication of (4) is the characterization of an event – that of the speaker producing the referring utterance – whereas the ‘ref’-predication is the characterization of a state. This is because texts refer because of the way they are, and not because of something they do.

(ii) The anchor set $\mathcal{K}_\tau$ of the ER for the text τ is assumed to contain a perceptual anchor that witnesses the interpreter’s current reading of it. But as often as not texts are known to us before we start reading them, and moreover, reading them often takes more than one session. So the interpreter’s ER for the text may have several anchors reflecting his previous encounters with it.

(iii) In contrast with this first ER of (8), I have assumed that the ER for the name token ν only has a single anchor, which witnesses the interpreter’s current perception of this token.

In the next section we will have a quick look at the static and their dynamic properties of coreference networks of the labeled representations of names. The discussion will be geared to the purpose of showing how such networks can be thought of as a kind of use-based denotations for the names that generate them, as the sets of those labeled ERs and pseudo-ERs that all represent the same referent of the same name.

3.7 ER Networks are Graphs

The structure of ER networks is a complex but important topic. This is not the place for going into any of the many intriguing features of such networks and I will limit myself to what is crucial in connection with the proposal in Section 4.

Formally, an ER network is best thought of as a directed graph, in which the nodes are ERs (and, in the case of N -labeled ERs, pseudo-ERs for tokens of N in texts) and the links established by vicarious anchors are the edges. We must distinguish, however, between ER networks existing at particular moments of time, consisting of ERs in the minds of the members of the speech community at that time and existing pseudo-ERs, and networks as they develop in the course of time – as sequences of ‘momentary’ networks at successive times in the history of the language. ER-networks that develop over time have both *static* properties – these are the properties of their momentary ‘time slices’, and properties which these time slices have just in virtue of what they are by themselves, without reference to other slices in the sequence – and *dynamic* properties, which have to do with how later slices develop out of earlier ones. It is most of all the dynamic aspects of ER-networks that makes them relevant for the purposes of this chapter.

Directly relevant to those purposes are parts of the over-all ER-network of a changing language community which consist of all N -labeled ERs, for some given name N , and the pseudo-ERs for the tokens of N . I will refer to such part networks as *N-networks*. But note that there is a potential ambiguity concealed in this terminology, which has to do with the identity conditions for names. Names can either be identified just in terms of their phonological-orthographic form, or in terms of their forms together with their referents. The latter option is often adopted within philosophy. My own preference goes towards the first option, for reasons that cannot be discussed here (but see Kamp (2015)). When names are identified in this way, however, then N -labeled ERs will in general not be coreferential, since they may be connected to different referents that go by the name N . So, simplifying somewhat, the totality of N -labeled ERs will be partitioned into disjoint subsets of mutually coreferential ERs.¹⁵

¹⁵The simplification has to do with the possibility of mishaps in the transfer of information using N : a speaker or writer may use N to refer to one thing and the recipient may take her to have referred to another thing that also goes by the name N . See the discussion on pp. 22, 23.

When names are identified purely in terms of their form, then referential ambiguity of names is ubiquitous. Nevertheless I will ignore this aspect of our use of names. Doing so will greatly simplify the discussion, and for our purposes there will be no serious loss of generality. So from now on it will be assumed that for any (non-fictional or fictional) name N , all N -labeled ERs will be coreferential; there is just one such set of mutually coreferential ERs.

With these simplifications it is possible to adopt some quite simple bits of formal notation. Let L be the language under discussion. For any given name N used in L and any time t in the history of L with its speakers and texts, the momentary N -labeled ER network at t will be denoted as $[[ER]]_{N,t}$. $[[ER]]_N$ denotes the N -labeled ER network as it develops through time – that function which gives the momentary network at t for any time t in L 's history.

3.8 Communicating about Entities in their Absence

A salient feature of the story I have told about the role that names play in linguistic communication is that their referents play no direct role in these transactions. This of course isn't surprising, nor is it something that is restricted to our use of names. One of the main points of language is that it enables us to communicate about what is not within sight or hearing, and thus isn't there to be pointed at or presented in any other non-linguistic way. But nowhere is this aspect of language use more plainly and dramatically manifest than in our use of proper names. The very point of having names is that we can talk concisely about their referents in situations where these referents are absent, and therefore not accessible by any other than linguistic means. And this does not only apply to entities that just happen to be absent on the occasion that we refer to them by their name, but also those that cannot be there because they no longer exist. It is especially the use of names for past entities that gives us a lead into the discussion of fictional names in Section 4.

How do I learn more about an entity $\mathbf{d}$ for which I have established an N -labeled ER? That depends on my relations to $\mathbf{d}$. If I can observe $\mathbf{d}$ directly, then I can learn much from those observations and for entities with which I interact on a regular basis this may be my main source of information about them. But for past entities this is not so. I have no direct access to them and my means of acquiring more information about them is much more restricted. By and large it is restricted to other information sources – to what I can learn

from texts and what other people can tell me. In fact, this will have been the way in which I came to adopt my N -labeled ER for $\mathbf{d}$ in the first place. In the terms we have adopted: All events that led to the introduction of my ER ER_N and to the information I collect about the properties of the entity it represents add vicarious anchors to ER_N . Vicarious anchors are all that can be found in its anchor set.¹⁶

Acquisition of information about a past entity $\mathbf{d}$ that is going by the name N is, then, a matter of transfers from one information source to another, and involving coreferential N -labeled ERs (or pseudo-ERs associated with tokens of N) that belong to the respective sources. Each such transfer produces a new link between two such ERs, thereby extending or reinforcing the network $[[ER]]_N$ of N -labeled ERs. This network can be thought of as the community's collective memory of $\mathbf{d}$ and all transfers of information about $\mathbf{d}$ can be described as transfers of information about this item of collective memory, without any direct reference to the entity $\mathbf{d}$ of which it is the collective memory. Formally this redescription would replace predications of $\mathbf{d}$ by predications of $[[ER]]_N$. For instance, if $\mathbf{d}$ is a historical figure and P is the predicate of being French, then the predication ' $P(N)$ ' will now be interpreted as predicating being French of $[[ER]]_N$.

Along these lines it is possible to work out an interpretation of repre-

¹⁶Some qualification of this may be needed. Many past entities – and, I guess, nearly all for which we still have names – have left traces of some kind or another. For instance, when $\mathbf{d}$ is a person, there may be authenticated portraits. Suppose I look at such a portrait while being told, truthfully, that this is N . What kind of contact is that and what kind of anchor does it give rise to? Is this an occasion where I am *perceiving* $\mathbf{d}$, and that gives rise to some kind of perceptual anchor; or is the portrait more like a kind of text, which determines a pseudo-ER for $\mathbf{d}$ and gives rise to a vicarious anchor of sorts? The question is not easy to decide, and when we look more closely we find ourselves on one of those slippery slopes that have been the downfall of many a philosophical claim. Let me say just this about the question here. Even if the conclusion is reached that portraits and other traces of past entities afford a fundamentally different access from sources like texts and human testimonies, this is nevertheless a significantly different kind of access from the direct and often interactional access that is possible with entities in our current environment. For one thing the information that I can learn about $\mathbf{d}$ by looking at a portrait is strongly limited – to what $\mathbf{d}$ looked like or perhaps what kind of clothes or jewelry $\mathbf{d}$ wore. Contact with other kinds of traces may yield other kinds of information. But generally the information that can be gathered in this way falls well short of what we can learn from observing the entities with which we interact directly. I am setting the matter of contact with past entities via their traces aside and operate on the simplifying assumption that the anchor sets of ERs for past entities only contain vicarious anchors.

sentations in language and thought – in MSDRT this would be some DRS language – as about ER networks in lieu of the entities represented by them.¹⁷ Note well, I am not recommending such a reinterpretation as a serious candidate for the *semantics* of the representations in question. The intentionality of these representations targets the entities represented in ER networks and not the networks as structures in their own right. It is the entities themselves, with their real world properties, that determine the truth conditions of the representations.

But the possibility of reinterpreting mental and linguistic representations along the indicated lines contains an important moral no less. ERs and ER networks are part of a toolkit for describing many aspects of how information is acquired, cognitively processed and disseminated and the role that names play in this. As far as these aspects of cognition and communication are concerned there is little difference between the non-fictional names considered so far and the fictional names that will be discussed in Section 4. The crucial differences between fictional and non-fictional discourse, and fictional and non-fictional names in particular, is the question of truth. For non-fictional discourse an interpretation theory of the type alluded to in the last paragraph is, by my books, nothing more than some sort of pseudo-semantics.¹⁸ But for fictional discourse a semantics of this type is all we can hope for.

4 Names in Fiction

4.1 Representing Fictional Content as Part of Mental States

How do we represent the contents of fictional texts? The basic assumptions I will make about this follow the lead of the proposal made by Maier in Maier (2017). Maier adopts a version of Walton’s (1990) principle that works of fiction (and works of art more generally) are invitations to imagine. For fictional texts the imagination must take its point of departure from the content that the text expresses by virtue of the grammar (the syntax and semantics) of the language in which it is written. It is an interesting question

¹⁷There is no room here to go into the details of such a reinterpretation. But the details are straightforward and of little independent interest.

¹⁸For someone with different metaphysical inclinations it might serve as a framework for a coherence theory of truth. See Young (2018)

what forms the imagination may take that fictional texts invite their readers to engage in. But I believe that by and large there is such a thing as a basic understanding of discourse that fictional texts share with non-fictional ones; and that a text must be understood – must have been semantically processed – at that level before the powers of the imagination can come into action. This chapter assumes that there is such a distinction and it will be only concerned with the results of semantic processing of texts at the first, basic level. But the format I will propose for the interpretations of fiction is designed to leave room for further interpretational steps, in which the imagination can take proper flight.¹⁹

Maier’s proposal is couched in his ADT (‘Attitude Description Theory’), the semantic formalism mentioned in Section 2. The use I will be making of MSDRT in this section is limited to its ADF part (also described in Section 2). This means that as far as the representation format for mental states is concerned, the two approaches are closely compatible.²⁰ Among the things both assumed by Maier’s proposal and the one I am making here about the representations that agents build for the fictional texts they read is that these representations occupy parts of the reader’s mental state that are separated from the state’s further contents. In fact, each work of fiction that the reader has read has its own representation compartment. Sometimes there will be cross connections between compartments. An example would be the compartment for Ovid’s *Metamorphoses* (with its many sub-compartments for the different stories that the *Metamorphoses* tell). The stories from the *Metamorphoses* are (or at least are presented as) stories from Greek Mythology. If the reader of the *Metamorphoses* has a Greek Mythology compartment, and is aware when she is reading Ovid that she is reading stories from Greek

¹⁹There are fictional texts – sometimes referred to as ‘experimental’ – which play loose and fast with grammar and where a clear separation between basic semantic processing and imagination of the kind I want to assume here does not seem meaningful. (Dadaistic poetry is of this sort.) This chapter has nothing to say about texts that do not abide by the grammar that is shared by all competent speakers of the language. It might be contested that the distinction between ‘experimental’ and ‘non-experimental’ texts isn’t a sharp one and that the distinction between a basic interpretation of a (non-experimental) text and more vivid or fanciful construals that use it as their point of departure is never really possible. This too is a question I won’t have anything to say about here.

²⁰There are some differences between ADF and the corresponding part of ADT, for instance in the forms that Maier and I assume for ERs. Although these differences are of some importance for what follows, I won’t say more about them, leaving the comparison for possible debate at some later time.

Mythology, then that awareness must result in close connections between these two compartments. Such cross-connections between works of fiction and fictional traditions like mythologies and folk tales are yet another aspect of fiction about which this chapter won't have anything to say.

4.2 Taking for granted: A modification of ADF and MSDRT

There is one feature that ADT (in the presentations of it that I know) and ADF as presented in Section 2 have in common and that I want to modify before we proceed further. It is a modification that affects in the first instance the representation of non-fictional information and only in the second instance, by a kind of analogy, the representation of fiction. In fact, I could have introduced this modification earlier. But it was thinking about the representation of fiction that led to it, and that is the reason, for what it is worth, why I have left this change until now.

According to the mental state descriptions we have been using so far all information that an agent considers true is represented in the form of beliefs. That isn't altogether wrong – surely what you hold true you believe and what you believe you hold true. But representing everything that plays the part of true information for the agent in the form of explicit beliefs is nevertheless missing a distinction of cognitive importance. This is the difference between a propositional content that you perceive as challengeable but consider true because you take yourself to have sufficient evidence for it, and information that you take for granted and where the possibility that it might be wrong won't play any part in your cognitive processes so long as nothing has come up to challenge it.

The distinction between information that one takes for granted, often without even being aware that one is doing so, and information that one consciously subscribes to as a belief for which one takes oneself to have adequate support, and that one is prepared to defend against potential skeptics, is revealed by exchanges of the form shown in (9)²¹.

- (9) *A*: Where the hell are my car keys? I have been looking everywhere.
 B: Have you looked in your car?

²¹Examples of this kind have been discussed extensively over the past 15 years or so. See for instance Franke and de Jager (2011).

A: My God, I didn't think of that. How stupid of me.

The point of this example is that *A* has simply not thought of his car as a possible place where the keys could be. That it is a possible place simply hadn't occurred to him. He realizes that it is a possible place only when *B* raises this possibility. Up to this point *A* has taken the assumption that the keys are not in his car implicitly for granted, and has been acting on this implicit assumption. But when *B* puts her question to him, he becomes aware that this is one of the assumptions that have been guiding his actions. (9) is different from the general case considered above in that *A* doesn't just turn a thus far implicit assumption into an explicit belief – in that case he would have reacted to *B*'s question with something like 'I hadn't really thought of that possibility. But no, I am pretty sure that they aren't there.' But common between the two cases is that an outside influence causes the agent to turn something that was only implicit in how he reasoned and acted into the content of a conscious attitude.

In the modification of ADF I am proposing *A*'s mental state before *B*'s question would be one which does not contain an attitude of the form $\langle \text{BEL}, \text{'The keys are not in the car'} \rangle$, but which excludes this possibility even so, in that it determines a doxastic background that consists exclusively of worlds in which the keys are not in the car. The effect of *B*'s reaction to *A*'s exclamation is that *A* now comes to acknowledge this possibility, thereby widening his doxastic horizon so to speak. At this point the proposition that the keys aren't in the car, which he had been taking for granted, is syphoned off from the totality of taken-for-granted information and turned into the content of an explicit attitude.

To capture the distinction between information that is taken for granted and challengeable beliefs, we admit, besides the types of constituents of mental state descriptions that were mentioned so far, also DRSs unaccompanied by a Mode Indicator. Each of these DRSs represents information that is taken for granted by the possessor of the described state. (10) is an example of a mental state description of this new kind. It describes the mental state of some agent *A* who has entity representations for Paris, London and a certain person called Mary, a cousin of hers, and who believes that Mary is in Paris but would like her to be in London instead. The two DRSs at the bottom of (10) represent two (unrelated) bits of information that *A* takes for granted: (i) that the Mary represented by the Entity Representation displayed is her cousin, and (ii) a bit of general knowledge, to the effect that water boils

when sufficiently heated. (The choice of this second bit is arbitrary and only meant as a hint at the wide variety of different kinds of information that we carry with us as background assumptions and that govern our thoughts and actions.) (i) is an example of information that the agent takes for granted about someone for whom she has an entity representation.

$$(10) \left\{ \begin{array}{l} \left\langle [ENT, m], \begin{array}{|c|} \hline \\ \hline N'd(m, Mary) \\ \hline \end{array}, \mathcal{K}_m \right\rangle \quad \left\langle [ENT, p], \begin{array}{|c|} \hline \\ \hline N'd(p, Paris) \\ \hline \end{array}, \mathcal{K}_p \right\rangle \\ \\ \left\langle [ENT, l], \begin{array}{|c|} \hline \\ \hline N'd(l, London) \\ \hline \end{array}, \mathcal{K}_l \right\rangle \quad \left\langle BEL, \begin{array}{|c|} \hline s \\ \hline n \subseteq s \\ s: in'(m, p) \\ \hline \end{array} \right\rangle \\ \\ \left\langle DES, \begin{array}{|c|} \hline s' \\ \hline n \subseteq s' \\ s': in'(m, l) \\ \hline \end{array} \right\rangle \\ \\ \begin{array}{|c|} \hline \\ \hline cousin'(m, i) \\ \hline \end{array} \quad \begin{array}{|c|} \hline \\ \hline \text{"water boils when} \\ \text{heated (enough)} \\ \hline \end{array} \end{array} \right\}$$

But is this always right way of formalizing taken-for-granted information? What exactly does it mean for an agent to take certain information for granted? The discussion above speaks of ‘implicit information’, information that reveals itself through the agent’s behavior. There is no compelling reason why such information should be present in the agent’s mind in the form of an explicit content representation. It might be present in such a form, and in that case its representation could enter into the processes that lead to the agent’s decisions and actions along some such lines as are assumed in formal accounts of practical reasoning. But for all that is currently known about the workings of the human mind the way in which implicit assumptions influence our decisions and actions could also take quite different forms, in which no explicit content representations are involved. On this point it is best, it seems to me, to remain agnostic so long as not more is known about the different

ways in which information can be internally available to the human mind and have its impact on its mental processes.

Of course this doesn't mean that information which an agent takes for granted is never explicitly represented. Plausible candidates are propositions about the physical world that in practice we tend to rely on without thinking, such as that water comes to the boil when you heat it enough, or even the more specific scientific law that the boiling point of water is 100 degrees Celsius. But so long as not all content that is taken for granted has an explicit representation, then what is represented explicitly will fall short of fully determining the agent's *doxastic horizon*, where by the doxastic horizon of the agent we understand the set of those worlds which verify everything that the agent takes for granted, the explicitly represented information as well that for which she has no explicit representation.

To sum up this section on MSDRT modification: Two modifications have been proposed, a change in the format of ADF descriptions of mental states and a change to the identity conditions of mental states. The first change consists in admitting content representations without Mode Indicators, as representations of information that the agent takes for granted. The second change is that the identity of a mental state is now determined in part by its doxastic horizon. The doxastic horizon of the mental state of an agent is some subset of the set of all worlds that verify all the beliefs in the state and all the explicitly represented information that is being taken for granted. It will be a proper subset when some of the information that is taken for granted does not have an explicit representation.²²

4.3 Back to the Representation of Fiction

As noted in Section 4.1, we will assume that the content representation which results when an agent A reads a fictional text T occupies its own compartment of A 's mental state, and that this compartment is marked as dedicated to the representation of T . First some words about the form of this dedication marking. When in Section 3.4 we discussed the interpretation of the occurrences of names in non-fictional texts we made the assumption that the interpreter has an ER for the text. I will make this same assumption also in

²²In the explicit formalization of MSDRT the second change will only be manifest in the model theory for its semantic representation languages. In what follows here, in which I won't make use of this formalization, we will assume doxastic horizons at an informal level.

relation to fictional texts. In addition we will make the default assumption that the reader has an ER for the author of the text.²³ This may not hold in all cases. Sometimes a reader may read a text without losing any thoughts over who might have produced it.²⁴ But normally the reader will have an Entity Representation for the author of the work as well as for the work itself. And as a rule that ER will be labeled with the author's name. (11) shows this pair of ERs for Tolkien's 'The Lord of the Rings'.²⁵

$$(11) \left\{ \begin{array}{l} \left\langle [ENT, tk], \begin{array}{|c|} \hline \\ \hline \text{Named}(tk, Tolkien) \\ \hline \end{array}, \mathcal{K}_{tk} \right\rangle \\ \left\langle [ENT, lr], \begin{array}{|c|} \hline \\ \hline \text{Named}(lr, TheL'd.o.t.R's) \\ \hline \text{by}(lr, tk) \end{array}, \left\{ \begin{array}{|c|} \hline s \\ \hline n \subseteq s \\ s: \text{read}(i, lr) \\ \hline \end{array} \right\} \right\rangle \end{array} \right\}$$

The next thing we need is a representation format for the part of the reader's mental state which contains her semantic representation of the text. I will stick fairly closely to the format that has already been used in MSDRT for the representation of propositional attitudes, as pairs whose first member is

²³Recall the decision of Section 3.6 to set aside cases of multiple authorship.

²⁴Also, there are texts for which the question of authorship is problematic: is it possible to assign authorship in their case? An example are Grimm's fairy tales. The tales appearing in this collection were written down by the Grimm brothers. But they are folktales the Grimms collected and which they then (re-)told in a form they considered suitable. There clearly is a sense in which the Grimms are the authors of these stories, But there is also another sense, and perhaps a more important one, in which they are not. (Ovid's *Metamorphoses*, mentioned earlier in Section 4, raise a similar issue.) Like the connections between the *Metamorphoses* and Greek Mythology, the questions raised by stories picked up and retold are among those I ignore in this chapter. We confine ourselves to literary works that come entirely from a single author, who is responsible both for the plot of the story and for the way it is told.

²⁵The ER for the text 'The Lord of the Rings' in (11) has only a single perceptual anchor (with 'read' as the perceptual predicate), which specifies the referent as currently being read by the agent. Admittedly this isn't very realistic. The most common for the reader of a novel or other fictional text is that she knew about the book before reading it and thus that she had an ER for it that goes back to the time when she first learned about the text. In that case her ER for the the text at a time at which she is reading it or after she has finished reading will have additional anchors, among them vicarious anchors in case she first heard about the text from others.

something like an attitudinal Mode Indicator and whose second member is a content representation. However, the basic content representations are now more complex than those of the propositional attitudes discussed in Sections 2 and 3. They are like the mental state descriptions of the modified version of MSDRT introduced in the last section.

To mark the compartment dedicated to T itself we make use of an indexed family of Mode Indicators as shown in (12).

(12) $\text{IMAG}(\tau)$

Here the parameter τ is the distinguished dref of the agent's ER for the text T . I am calling the marking devices in (12) Mode Indicators because they indicate the cognitive status of the parts that they mark, viz. as the agent's representation of the text represented by her ER with distinguished dref τ , just as, say, the Mode Indicator BEL marks the content representation following it as representing the content of a belief. The expression 'IMAG' has been chosen as a tribute to Walton's idea that fictional texts are invitations to the reader to imagine a world in which their words come to life.

A more difficult question is what may go into the compartments marked by the Mode Indicators in (12). This is a question to which I neither can nor want to give a complete answer at the present time. But let me say this much: At a minimum what should be allowed to go into a text dedicated compartment are representations with as much structure as the mental state descriptions spoken of hitherto. Such representations may have (i) a taken-for-granted-component, as specified in the ADF representations of Section 4.2, which gives the 'basic level' semantic content representation of the text, (ii) ERs for the fictional characters that the text talks about and whose names appear in it and (iii) various propositional attitudes, which represent thoughts that the reader may have about the content of the text as she is reading it – thoughts that are suggested by the text to her imagination, among them opinions that she may form about some of its characters or feelings that she develops towards them. But texts and their interpretations have much more structure than this. To deal with the many structural properties of fictional texts that narratology and other branches of literary science have identified and studied further representational devices will have to be put in place. (For instance, suitable devices will be needed for the representation of everything that has to do with the role of the narrator or narrators of a text, as distinct from its author). From the perspective of literary studies the

contents of text dedicated compartments is no more than a very first step, and one from which those engaged in such studies will find missing most of what they are primarily concerned with.

(13) gives an example of a mental state description which has a compartment dedicated to the content of Tolkien's 'Lord of the Rings'. (13a) shows the description over-all and (13b) more details of the 'Lord of the Rings' compartment. (The symbol ' $\mathcal{R}$ ' in (13a) is a shorthand for the structure shown in (13b).) (13a) is an extension of the merge of the mental state descriptions in (10) and (11) with in addition the ERs for Tolkien and 'Lord of the Rings' and the Condition ' $\langle \text{IMAG}(lr), \mathcal{R} \rangle$ '.

$$(13) \quad \text{a.} \quad \left\{ \begin{array}{l} \left\langle [ENT, m], \frac{\quad}{\text{N'd}(m, Mary)}, \mathcal{K}_m \right\rangle \quad \left\langle [ENT, p], \frac{\quad}{\text{N'd}(p, Paris)}, \mathcal{K}_p \right\rangle \\ \left\langle [ENT, l], \frac{\quad}{\text{N'd}(l, London)}, \mathcal{K}_l \right\rangle \quad \left\langle [ENT, tk], \frac{\quad}{\text{N'd}(tk, Tolkien)}, \mathcal{K}_{tk} \right\rangle \\ \left\langle [ENT, lr], \frac{\quad}{\text{N'd}(lr, Lord.o.t.Rings)}, \left\{ \frac{s}{n \subseteq s} \right\} \right\rangle \\ \quad \frac{\text{by}(lr, tk)}{\quad} \\ \left\langle BEL, \frac{s}{n \subseteq s} \right\rangle \quad \left\langle DES, \frac{s'}{n \subseteq s'} \right\rangle \\ \quad \frac{s: \text{in}'(m, p)}{\quad} \quad \frac{s': \text{in}'(m, l)}{\quad} \\ \frac{\quad}{\text{cousin}'(m, i)} \quad \frac{\quad}{\text{"water boils when heated (enough)"}} \\ \langle \text{IMAG}(lr), \mathcal{R} \rangle \end{array} \right\}$$

$$\text{b.} \left\{ \begin{array}{l} \left\langle [ENT, f], \frac{\boxed{}}{N'd(f, Frodo)}, \{\emptyset\} \right\rangle \left\langle [ENT, b], \frac{\boxed{}}{N'd(b, Bilbo)}, \{\emptyset\} \right\rangle \\ \left\langle [ENT, r], \frac{\boxed{}}{N'd(r, the Ring)}, \{\emptyset\} \right\rangle \left\langle [ENT, d], \frac{\boxed{}}{N'd(d, Mt. Doom)}, \{\emptyset\} \right\rangle \\ \left[\begin{array}{c} e_1 \quad e_2 \\ \hline e_1 \prec n \quad e_2 \prec n \\ e_1: \text{get-from}'(f, r, b) \\ e_2: \text{carry-to}'(f, r, d) \end{array} \right] \left\langle JUDG, \frac{\boxed{\text{brave}'(f)}}{'(f)} \right\rangle \end{array} \right\}$$

I will refer to $IMAG(\tau)$ -marked parts of mental state descriptions like (13a) as their T – *dedicated compartments* (where T is the text represented by a text-representing ER with distinguished $dref \tau$). (So (13b) is the compartment in (13a) that is dedicated to the text ‘The Lord of the Rings’.) (13b) shows besides a (vanishingly small) part of the basic semantic representation of ‘The Lord of the Rings’ also one constituent that has the form of a propositional attitude. The Mode Indicator JUDG I have chosen for this constituent is meant to capture the notion that evaluative judgments of individuals constitute their own category of propositional attitudes, which requires its own form of semantic evaluation.²⁶

When its Judgment constituent is ignored, (13b) gives a minute fragment of the representation of ‘The Lord of the Rings’, consisting of four ERs and a couple of propositions involving them. Evidently there is no way of evaluating this representation for truth or falsity in the way this is possible for non-fictional representations. That is so for one thing because the ERs of (13b) which occur in the Conditions of the representation of this content do not

²⁶The question whether and in what sense various types of evaluative judgments have propositional content is a vexed one. Is there any sense for instance to the question whether or not Frodo was brave? Perhaps there is in this case, but the truth conditions are loose and easily contestable. There is an extensive literature on this topic. For insightful discussion see ?. This is one of the issues I am setting aside in this chapter.

have referents in the real world.²⁷ This is as it should be. ‘The Lord of the Rings’ isn’t about what actually happened in the world in which it was written and that is inhabited by the one who wrote and those who read the book. So these ERs cannot have anchors of the kind discussed in Section 3.2, which tie their ERs referentially to the world of the agent whose mental state they belong to.

As (13b) shows, the content representation of a T -dedicated compartment resembles the structure we have assumed for mental state representations over all. (So there is a kind of fractal dimension to the structure of mental state representations.) I will refer to the content representation of a T -dedicated compartment as $\mathcal{K}_T$ and to the T -representing DRS of $\mathcal{K}_T$ simply as K_T .

There is a further similarity that we will assume between T -dedicated compartments of mental states and mental states over-all. In Section 4.2 we modified MSDRT’s general notion of a mental state to the effect that a mental state is characterized in part by its doxastic horizon, a constraint on the set of possible worlds compatible with the representational structure of the state, in the sense that the worlds in this set verify all representations in the state of beliefs and all explicit representations of information that the agent takes for granted. Analogously I propose that the representation which an agent forms for a fictional text she reads will also come with a doxastic horizon, which is appropriate to that text. For many fictional texts this doxastic horizon will be the same as the agent’s doxastic horizon over all. These are the texts that do not require their readers to adjust their assumptions of what is possible in principle. But there are also texts – fairy tales, mythologies or science fiction – that force us to suspend some of our general assumptions about what is possible. (‘The Lord of the Rings’ is an obvious example.) Mostly when the doxastic horizon of a reader’s interpretation, DH_T , differs from her ‘real world related’ doxastic horizon DH_{RW} , DH_T results from discarding some of the assumptions that shape DH_{RW} (for instance, the assumption that there are no hobbits and no elves and no creatures or contraptions that the reader has never thought but would consider impossible if they were described to her). In such cases the world set DH_T will be a proper superset of the set DH_{RW} . Perhaps there are also pieces of fiction that impose new constraints of their own, so that this inclusion relation will not hold. But as things stand I am

²⁷This is indicated by the anchor sets $\{\emptyset\}$, with the single ‘dummy anchor’ $\emptyset$, which does not provide a link with anything. We will return to the question of anchoring links for ERs labeled by fictional names in the next section.

not sure of this.

4.4 Fiction Protagonists vs. Fictional Characters

4.4.1 The double life of Entity Representations labeled with fictional names

Fictional texts, Walton says, are invitations to imagine; they invite us to build a world – the world of the story – by exercising our imagination. Of course, the building of a story world will always and inevitably be incomplete, even when we let our imagination roam to the full of its powers. What gets built is always something like a blueprint for a world. And a blueprint is something that can be turned into a complete world in any one of a large number of different ways.²⁸

With each piece of fiction T , then, comes a range of possible worlds which all have equally good claims to being ‘the world of the story’. I take it as obvious that each of these worlds should (a) verify the reader’s text representation K_T and (b) belong the agent’s doxastic horizon DH_T for T . But speaking of K_T being true in a world w comes with a presupposition: the ERs that are part of $\mathcal{K}_T$ (and whose distinguished discourse referents will occur in K_T) must have referents in w . I see no other way to satisfy this presupposition than by stipulation: some of the worlds in DH_T are worlds of the kind T describes, and these are worlds in which the ERs of $\mathcal{K}_T$ all do have referents; and in some of these worlds K_T will be true, while in other such worlds K_T may be false.²⁹ To capture this intuition, let us assume –

²⁸From this perspective the expression ‘*the world of the story*’ is really a misnomer. Yet it is curious how natural the phrase feels, even though everyone is aware that a story cannot (and shouldn’t) tell you everything about ‘its’ world. But for the reader who lives through the events of the story as her reading progresses the plurality of story worlds plays no part in her imaginative experience. The story world that unfolds before her inner eye and to which her own imagination makes its contributions isn’t all that different from what goes on when one is trying to learn more about an episode that actually happened. When compared with the totality of all that could be learned the information one has gained is always no more than a fraction; there is always more to find out. Often that includes things that we would like to know, though there will also be no end of things that would be of no interest. In fact, whatever the phenomenological differences may be between the experience of building a world of fiction in our head when reading a fictional text and the experience of learning more about a situation or episode in the real world, completeness of information does not seem the decisive factor.

²⁹Unless the text T is inconsistent or tautologous. I set these marginal possibilities aside.

this is the stipulation I am making! – that with each fiction interpretation $\langle \mathcal{K}_T, DH_T \rangle$ comes a 2-place function REF from the ERs in $\mathcal{K}_T$ and the worlds w in DH_T which maps pairs $\langle ER, w \rangle$ to entities in w . Intuitively $REF(ER, w)$ is a referent for ER in w . So REF can be thought of as filling (by stipulation) the gap left by the vacuous anchor sets of the ERs that readers of T introduce for its protagonists.

In the discussions of the remainder of Section 4 I will assume that for each interpretation of a text T a function REF_T is given. So an agent's interpretation of a text T , as captured by the T -dedicated compartment of her mental state, will henceforth be assumed to have three components, $\mathcal{K}_T$, DH_T and REF_T . Observe that for any such interpretation $\langle \mathcal{K}_T, DH_T, REF_T \rangle$ the world set DH_T is divided into two parts, one part consisting of the worlds w such that the content representation of T is verified in w when the distinguished drefs of ERs that occur in it are assigned the REF values in w of those ERs, and the other part consisting of the remaining worlds. We will refer to the first part as $DH_{T,tr}$, the truth set of the interpretation.

The function REF provides the ERs in an agent's T -dedicated compartment with some kind of 'referents within the story'. To see this, note that in the worlds of $DH_{T,tr}$ – those in which the content representation K_T is true – REF_T will map the ERs of $\mathcal{K}_T$ to the kinds of entities they are according to T . This will be so provided that K_T represents the relevant sortal information about the protagonists of T . For an illustration, assume T once more to be 'The Lord of the Rings'. It is hard to imagine that the K_T of any reader of this story lacks a Condition of the form 'hobbit'(f)' where f is the distinguished dref of the ER ER_{Frodo} that she has used in her interpretation of the occurrences of the name *Frodo* in T . The verification of K_T in w entails the verification of this Condition; and that can be the case only if $REF(ER_{Frodo}, w)$ is a hobbit.

But there is more that readers of fiction do with what they get out of the texts they read than building their own private imaginary worlds. We often talk about fictional texts with others, write comments about those texts, compare them with other texts, draw parallels between them and the actual world in which we live. I will refer to all such activities and to the written and spoken utterances involved in them and the thoughts we thereby express as *meta-fictional*. Fictional names play a prominent part in these metafictional activities. But these metafictional uses are different from the uses they play when we first read a fictional text and turn its names into protagonists of the world our imagination builds.- And it is these different, metafictional uses

of fictional names that, I believe, we are most naturally disposed to see as referring to ‘fictional characters’. But then the question is: What are these?

Before I propose an answer to this question, there is another proposal that needs to be discussed first. This is that ERs belonging to fiction-dedicated compartments of mental states can be *exported* to other parts of the agent’s mental state; for each compartment-internal ER it is possible to create a compartment-external duplicate. It is these external duplicates that figure in meta-fictional thought and talk and it is they, I propose, that represent fictional characters (as opposed to the compartment-internal counterparts of which they are the exported duplicates, which represent non-existing entities of other sorts, like hobbits or mountains).

There are close connections between compartment-internal ERs ER and their exported duplicates ER' . We will capture these connections through their labels: When ER is labeled with the name N , then so is its duplicate ER' . We also assume that name-labeling takes the same form for fictional as for non-fictional names, as described in Section 3.5): ER is labeled by N when its descriptive component contains the Condition ‘Named(α, N)’, where α is the distinguished dref of the ER . But there are also important differences. First, the distinguished drefs of ER and ER' are always distinct. Secondly, and more importantly, external duplicates ER' represent entities belonging to the real world, whereas compartment-internal ERs labeled with fictional names never do. Finally ER and ER' differ in their anchor sets. The anchor set of compartment-internal ERs is assumed to always consist of a single ‘dummy anchor’, a signal that there is no entity to which it is linked in the manner of ERs that represent entities in the real world. But this is not so for the exported duplicates of such ERs. The anchor set of a duplicate ER' is non-empty from the moment it is created. At that point it will contain a single anchor, one of a new type, which links ER' to the ER ER that it is the duplicate of. This ‘anchor’ is different from the anchors we have been considering so far. It is like the vicarious anchors of Section 3.1 in that it links its ER (i.e. the ER ER') with another ER (viz. the compartment-internal ER ER of which ER' is the duplicate). But it is unlike those vicarious anchors in that it doesn’t anchor its ER to the referent of the ER at the other end of the link it witnesses. For that ER, the compartment-internal ER ER , doesn’t represent a referent.

An illustration of ER exporting is shown in (14). (14) is the result of exporting the internal ER for Frodo in (13b). (So (14) is to be seen as an extension of (13a); but some parts of (13a) have been dropped, in order to

make the over-all structure of mental state representations with exported duplicates more easy to perceive.)

$$(14) \quad \left\{ \begin{array}{l} \left\langle [ENT, m], \frac{}{N'd(m, Mary)}, \mathcal{K}_m \right\rangle \quad \left\langle [ENT, p], \frac{}{N'd(p, Paris)}, \mathcal{K}_p \right\rangle \\ \left\langle [ENT, l], \frac{}{N'd(l, London)}, \mathcal{K}_l \right\rangle \\ \left\langle [ENT, tk], \frac{}{N'd(tk, Tolkien)}, \mathcal{K}_{tk} \right\rangle \quad \left\langle [ENT, lr], \frac{}{N'd(lr, LotRs)}, \left\{ \frac{s}{n \subseteq s} \right\} \right\rangle \\ \left\langle [ENT, f'], \frac{}{N'd(f', Frodo)}, \{\mathbf{1}_f\} \right\rangle \\ \left\langle BEL, \frac{s}{n \subseteq s} \right\rangle \quad \left\langle DES, \frac{s'}{n \subseteq s'} \right\rangle \quad \frac{}{cousin'(m, i)} \\ \left\langle IMAG(lr), \left\{ \begin{array}{l} \mathbf{1}_f : \left\langle [ENT, f], \frac{}{N'd(f, Frodo)}, \{\emptyset\} \right\rangle \quad \left\langle [ENT, b], \frac{}{N'd(b, Bilbo)}, \{\emptyset\} \right\rangle \\ \left\langle [ENT, r], \frac{}{N'd(r, the \ Ring)}, \{\emptyset\} \right\rangle \quad \left\langle [ENT, d], \frac{}{N'd(d, Mt.D.)}, \{\emptyset\} \right\rangle \\ \frac{e_1 \ e_2}{e_1 \prec n \quad e_2 \prec n} \\ e_1: \text{get-from}'(f, r, b) \\ e_2: \text{carry-to}'(f, r, d) \end{array} \right\} \right\rangle \\ \left\langle JUDG, \frac{}{brave'(f)} \right\rangle \end{array} \right\}$$

The elements of (14) that are new (i.e. not in (13a)) have been bold-faced for easier spotting. They are: (i) the label $\mathbf{l}_f$ of the compartment-internal ER for Frodo and (ii) the Mode Indicator [ENT, f'] and the anchor of the external ER that is the result of exporting the compartment-internal ER for Frodo.³⁰

Whenever we think or talk about a piece of fiction, this will involve external duplicates of ERs that are internal to our representation of the fiction that our thoughts or talk are about. (Either the duplicate involved is already in place or it is created as part of conceiving the thought or of producing or interpreting the talk.) Moreover, once it has been put in place a duplicate can interact with the ERs of other agents in much the same way as this was outlined in Section 3 for ERs that are part of agents' representations of the real world. Through such interactions these ERs too are integrated into ER networks. And another way in which duplicate ERs resemble ERs that represent real world entities is that once established, they can become the targets of further thoughts.

All this applies in particular to duplicates of ERs that are labeled with fictional names. For instance, when we talk with others about some fictional text and make use of the names occurring in the text, then it will be our exported duplicates labeled by those names that are involved in our active and passive contributions to the discussion and that get linked more intimately to each other in the course of it. And by much the same token, external ERs labeled by fictional names can also be the cumulation points for further

³⁰The formal device used to represent the links that connect exported ERs with their compartment-internal counterparts is what in MSRT is called *internal labeling*. Internal labeling is used for various purposes in MSDRT, not just in the special application that is made of it here. The original motivation for adopting it was that when we reflect on our own thoughts – for instance, when we question whether we are right in sticking to a belief to which we have long been committed, or when we form the desire to get rid of a desire we have (for one last cigarette, or for the object of an unrequited love) – we can focus on particular parts of our mental state. MSDRT assumes that the ‘first order thoughts’ that are the targets of the ‘second order thoughts’ we form when we engage in self-reflection are directly accessible to those second order thoughts. Labeling is used as a formal device to capture this directness: The first order thought is given a label and the content representation of the second order thought uses this label to refer to the labeled thought (much in the way that the distinguished discourse referents of ERs can be used in the content representations of propositional attitudes).

In the use that is made of labeling here the labeled constituents are the internal ERs ER_{int} and their labels are employed in the anchors of their external ‘duplicates’. In fact, the anchor of the external duplicate ER_{ext} of the internal ER ER_{int} will consist just of ER_{int} 's label, and nothing more.

thoughts about what we think of as the bearers of those names. For instance – to take just two arbitrarily chosen examples – someone who has read *The Lord of the Rings* may speculate on how often Frodo was on the verge of giving up on his mission; or she may suddenly recall that the first time she heard about Frodo was when as a small child she was read bits from *The Lord of the Rings* by her uncle.

The exported duplicates of labeled ERs inherit their labels from the ERs they duplicate. But not all external ERs labeled with fictional names are the result of exportation. Some people participate in exchanges about fictional texts and make use of names from those texts without ever having read or listened to the texts themselves. They may pick up things about the text in the course of these exchanges and part of that will take the form of adopting ERs labeled by the fictional names used by their interlocutors with vicarious anchors that link them with the ERs from other users of the names. But if they were never properly exposed to the fiction itself – never took the opportunity to engage with it in that way that can lead to the fiction-dedicated compartments that result from imaginative engagement – then their ERs won’t be the results of exportation, for the simple reason that they have nothing to export from.³¹

A good deal more could be said about the different ways in which ERs labeled by fictional names can be formed and used. But what I have said will do for the main purpose of this discussion: ERs labeled with fictional names form ER networks by virtue of the same mechanisms as the ERs labeled with non-fictional names discussed in Section 3 and as a consequence these networks share many of their structural properties. Moreover, the two kinds of networks are also like each other in that their ERs are mutually coreferential, albeit ‘coreferential’ in not quite the same sense. The networks discussed in Section 3 are coreferential in the sense that any two ERs belonging to the network represent the same real world entity. But such networks are also coreferential in a different, ‘intentional’ sense: When the N -labeled ER ER_B of one agent B has a vicarious anchor that links it to the N -labeled ER ER_A

³¹There is one qualification to this. It often happens that someone learns about the existence of a piece of fiction and one or more of its protagonists first and only then sits down to reading the text, which results in her dedicated compartment for it. In this case the external ERs she already had that are labeled with some fictional name N from the text will be identified with the internal N -labeled ERs that are the direct result of the reading. In other words, they are turned into exported duplicates although they weren’t that to begin with.

of some other agent A , then this means that B *regards* his representation ER_B as representing the same entity that is represented by ER_A , irrespective of what further information B may associate with ER_B . To repeat, it is these commitments that make it possible for users of a name N to understand each other as speaking of the same things even though they may have little or no first hand information about what those things are. It is in this second ‘intentional’ sense that networks $[[ER]]_N$ of ERs labeled by some fictional name N are coreferential too, even if there is no obvious sense in which they are coreferential in the first sense.

A second difference between the networks of ERs labeled by fictional and networks of ERs labeled by non-fictional names is that the former contain ERs of two different sorts, compartment-internal ERs and compartment-external ones, with special links between compartment-internal ERs and their exportations, effected by anchors like that of the exported ER for Frodo in (14). There is no such distinction between the ERs found in networks of the second kind. All ERs labeled by non-fictional names are ‘external’ in the terminology of the present section.

But for our purpose the similarities between the two kinds of networks outweigh the differences. It is to this purpose that we now turn.

4.4.2 Entity Representation Networks as fictional characters and as denotations of fictional names

At last we have reached the point for the answer I want to propose to the question: What is a fictional character? But let me reformulate the question slightly: What are the fictional characters denoted by fictional names?³² After all I have said the answer to this last question will probably be expected: The denotation of a fictional name N is the corresponding network $[[ER]]_N$. The intuition behind the answer should be clear at this point. But let me give it once more. Evidently, fictional names do not have referents in the way that non-fictional names do. What a fictional name N is about is determined on the one hand by the fiction to which N belongs and on the other by the *communis opinio* of the language community. $[[ER]]_N$ captures both of these: the first through the connections of the compartment-internal ERs of some

³²I am using the verb *denote* here for lack of a better alternative. (*refer* and *represent* have been used extensively in this chapter, with quite specific technical connotations, which disqualifies them here.) The substance of the denotation relation will become clear in the next section, about the semantics of expressions containing fictional names.

of community members to the piece of fiction in which the name originates and the second through the roles that each ER in $[[ER]]_N$ plays in the mind of its member of the community.

Unfortunately, whatever the merits of this intuitive justification, the answer cannot be left in the simple form in which I have stated it. ER networks change in the course of time, and they do that in two ways. First, there are changes to the individual ERs of which the network is made up, through the addition of new anchors or through changes to their descriptive component. Second, the network as a whole is subject to change in that new ERs get added to it, as new members of the community are introduced to the name, while other ERs are lost, when community members take their ERs with them to their grave. But if networks change over time, what exactly are the ‘networks’ that are supposed to be the denotations of fictional names at some given time t and that should play the part of fictional characters?

As far as I can see, there are three possible answers to this that are worth considering.

- (i) $[[ER]]_{N,t}$, the ‘momentary’ N -labeled ER network at t ;
- (ii) the sequence consisting of the momentary networks $[[ER]]_{N,t'}$ from the time t_0 when N was introduced into the language (as part of the fiction to which it belongs) all the way up to and including t ; (formally: the function $\lambda t':t' \in [t_0,t].[[ER]]_{N,t'}$);
- (iii) this same function, but from the time t_0 till the end of N ’s history, the result of either its obliteration from the culture and its language or because language and culture have disappeared altogether.

Each of these three answers has what may be considered its pros and cons. The first two have the effect that the denotation of N changes as time goes on: $\lambda t':t' \in [t_0,t].[[ER]]_{N,t'}$ depends on t . That may seem an odd thing for the denotations of names, given the now widely accepted view that names have fixed denotations. The third answer doesn’t have this consequence: the function it proposes as denotation of N is the same for all t in the history of N . But it has another feature that may strike one as odd: at any time t at which N is part of language and culture the function has been instantiated only in part. What will happen to $[[ER]]_{N,t'}$ at later times t' than t depends on how the world will develop after t and at t that may still be largely undetermined.

I do not see either of these consequences as counting decisively against the

answer or answers that entail them. To substantiate this, a quite extensive discussion would be needed and that would distract from the concerns of this chapter.³³ Furthermore, I do not have a compelling argument in favor of any one of the three options. So for the time being I leave this matter open. For expository reasons, however, it will be convenient to make a concrete proposal. In this spirit, let us provisionally adopt answer (i).

- (15) a. The denotation $[N]_t$ of a fictional name N belonging to a language L with language community at time t is the network $[[ER]]_{N,t}$.
 b. *fictional characters* are the denotations of fictional names N .
 c. The fictional character denoted by N at t is also the *entity represented at t by any N -labeled ER* $[[ER]]_N$.

Definition (15) discharges the central promise of this chapter: a characterization of fictional characters as the denotations of fictional names, which explains both what fictional names have in common with non-fictional names and what makes the two fundamentally different: Common between them is that both non-fictional and fictional names determine networks of name-labeled Entity Representations in the mental states of members of the language community and pseudo-ERs associated with names in texts. The difference is that non-fictional proper names have referents, to which they have been assigned as ‘labels’ through some kind of baptismal act that requires the referent as participant. The ER network for such a name is a collective representation of this network. The ER networks for fictional names lack such

³³Here are some hints to what I see as involved in the choice between answers (i), (ii) and (iii). On the one hand fictional names aren’t ‘real names’: they don’t have referents by virtue of having been made their labels through some kind of ‘baptism’. That weakens the claim that their denotations should be impervious to change. If the denotation of a fictional name is to reflect the distributed opinion of the community as to what it stands for, and this distributed opinion changes with time, then that is reason enough for the conclusion that the name’s denotation is subject to change. As regards the objection against option (iii): The problem that part of the denotation of N may not yet be determined at any time t at which N is in use because of indeterminacy of the future, is an instance of a much more general problem, other instances of which do not seem to greatly worry many of us: Is indeterminacy of the future properties of an entity at time t a problem for its identity at t ? The sense in which future indeterminacy might constitute a threat to current identity is brought out by theories that analyze entities whose existence is bound in space and time as *spatio-temporal continuants*: sequences of ‘time slices’ held together by some kind of metaphysical kit. For discussion see e.g. Wiggins (1967, 2001).

a referential basis. All they have is their role in communication and thought. That makes ER networks for fictional names plausible as denotations for their names, and therewith also plausible candidates for the role of fictional characters.

However, postulating ER networks as denotations of fictional names would be no good if they could not be used in the formulation of a semantics for discourse containing such names. This is the topic of the next and last subsection of Section 4.

4.5 Fictional Character Semantics

To repeat the concluding words of the last section: It is in a semantics of discourse involving fictional names that the fictional characters that I have proposed as their denotations will have to prove their mettle. But what are we to expect from such a semantics? Should it be an account of the truth conditions of discourse containing such names, or should it take some altogether different form? There are two distinct questions here: (i) Is a truth-conditional semantics for fictional and/or fiction-related discourse desirable? and (ii) Is such a semantics possible? As regards the first question I may not be the right person to ask. I have no idea what form a semantic account of any kind of language could take that doesn't make truth conditions its central concern. This means that for me the question whether a truth-conditional semantics for fictional and fiction-related discourse is possible amounts to the same as the question whether such a semantics is possible at all. So a truth-conditional semantics for fictional and fiction-related discourse better *be* possible. What follows are some preliminary explorations of such a truth-conditional semantics. Our focus will be – naturally, in view of the central concern of this chapter – on the role of the denotations for fictional names proposed in (15) of the last section.

One challenge for a semantics of discourses containing fictional names is their heterogeneity. This is illustrated by the following set of examples, each of which contains the fictional name *Frodo*.

- (16) a. Frodo carried the Ring to Mount Doom.
- b. You mentioned Frodo to me last night.
- c. You said last night that Frodo was taller than Sam.
- d. At that point Gollum bites off Frodo's finger. I think that was too much of a *Deus ex Machina*. But I loved the end nonetheless.

The natural way to understand (16a) is as a statement about the story Tolkien has told; (16a) is true if what it says is something that is part of that story – something that it says explicitly or that is entailed by what it says explicitly. On this understanding of it (16a) is surely true; anyone who knows the story will readily confirm this. But how can what we have defined as the fictional character denoted by *Frodo* help us to make this prediction? Certainly not in the way in which the referent or denotation of a name or other singular term of the language is normally assumed to enter into the truth conditions of statements containing it. According to that way (16a) would be true iff there was an event *e* which jointly with the referent of *Frodo* and the referents of *the Ring* and *Mount Doom* would be in the extension of the predicate *carry to*. But if the referents of *Frodo*, *the Ring* and *Mount Doom* are networks of Entity Representations that readers of *The Lord of the Rings* (and others without first hand knowledge of the text) have formed to represent its protagonists, then it is hard to see how such a quadruple – consisting of an event and these three referents for the names – could be a member of the extension of *carry to*. To qualify as arguments of *carry* these referents are obviously of the wrong sort. But if the denotations we have adopted for the fictional names *Frodo*, *the Ring* and *Mount Doom* cannot be arguments of the relational predicate expressed by *carry*, what contribution can they make to the truth conditions of a sentence like (16a)?

Before we look more closely into this question it will be helpful to first have a look at (16b). But let us retain this from the little that has been said so far: The truth conditions of sentences like (16a) are *partial*. Such sentences are true when their content is entailed by the relevant piece of fiction, they are false when their content is refuted by it, and when their content is neither entailed nor refuted, they are neither true nor false.

Sentences like (16b) are bivalent: Either you did mention Frodo to me last night and then (16b) is true, or you didn't and then it is false. But how are these truth conditions determined and what is the part that the denotation of *Frodo* plays in this? For (16b) this seems straightforward. Whether (16b) is true or false is determined by the extension of the predicate *mention*. If we assume that verbs are descriptions of eventualities – a more or less standard assumption in semantics by now and one that we have followed implicitly in all content representations in this chapters – then the ditransitive verb *mention* is a 4-place predicate, which expresses a relation between an (i) event, (ii) an agent (the one who does the mentioning), (iii) a theme (the entity mentioned), and (iv) a recipient (the addressee of the utterance in

which the mention is made). For instance, when I mention a friend of mine to you, the 4-tuple $\langle e, i, v, y \rangle$, consisting of e (the action of mentioning), i (me), v (my friend), v and y (you), belongs to the extension of *mention*. And if I don't, it doesn't.

Mentioning something is an act that involves language – the use of an expression α that denotes the something. Often α will be a name, as it is in the case of (16b). But what are the sorts of things that can be mentioned? What sorts of things can fill the third argument slot of this verb? In particular, what is it that according to (16b) the addressee mentioned to the speaker by means of her use of *Frodo*? That question isn't easy to answer on the basis of pre-theoretic intuition. Perhaps the most natural untutored reaction to it would be: 'What is mentioned in (16b) is a fictional character.' But then the next question would be: 'What are fictional characters?' And at this point a theory is needed.

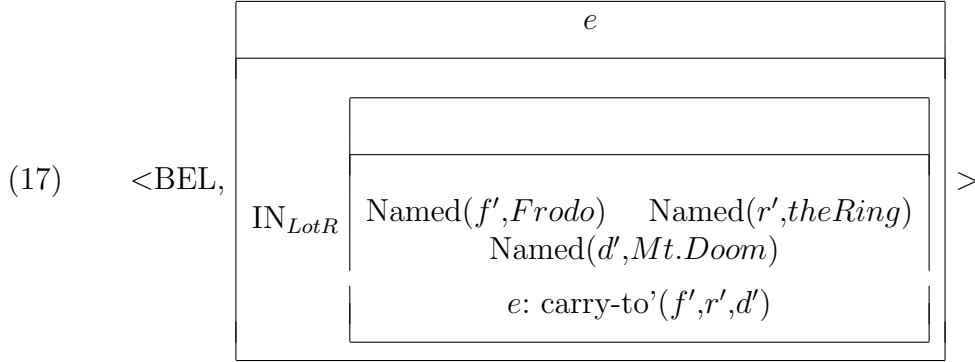
The upshot of this is that the semantics of *mention* is dependent on semantic theory in a way that the *carry* of (16a) is not. If it is part of the meaning of *mention* that its third argument slot can be filled by fictional characters, then the selection restrictions imposed on this argument slot depend of what our theory tells us fictional characters are. By telling us that it also tells us something about the semantics of the verb.

In this regard *mention* as it is used in (16b) is different from the verb *carry* as it is used in (16a). What sorts of entities can fill the argument slots of *carry* – entities like people or animals for the second slot, physical objects for the third slot, places for the fourth – is something we can figure out by reflecting on our lexical knowledge as speakers of the language. For *mention* this is not so. *mention* is a verb used to describe certain kinds of linguistic acts. It is a 'semantic verb', if you like, and that is why semantic theory is needed to settle what its semantics is.

In view of this there isn't much that the truth conditions of (16b) can tell us about the adequacy of the definition of fictional characters in (15). The truth-conditions should come out the same no matter how we define the denotations of fictional names. All that is demanded of the definition is that it assign fictional names *some* kind of denotations and that it assign different denotations to names for different protagonists. What matters for the truth of (16b) is whether there was an event of the speaker using the fictional name *Frodo* to mention the fictional character Frodo. Settling that question doesn't depend on any specific assumption about what fictional characters are – any specific assumption about their exact 'identity conditions'. Perhaps there is

a sense in which the *meaning* of (16b) depends on those identity conditions of fictional characters. If there is, then it is one to which the truth conditions of (16b) are insensitive.

After this look at the truth conditions of (16b) we return to (16a). What can we say about the contribution that the fictional name *Frodo* makes to the truth conditions of this sentence? As a first step consider what can be said about the *belief* that a member of the speech community could express by using (16a). More generally, let us consider what it is for the reader *A* of a fictional text *T* to entertain beliefs about what is the case according to *T*. There are two assumptions about such beliefs that I will make. The first is that it should be possible for these beliefs to be represented in *A*'s mental state outside its *T*-dedicated compartment; and the second is about how these compartment-external belief representations come about. Both assumptions are based on a third one: parts of the content representation of *A*'s *T*-dedicated compartment can be exported to locations outside this compartment, in much the same way as has been assumed already for compartment-internal ERs. More specifically, we will assume that a compartment-external belief representation of some part of the content of *T* will have the form $\langle \text{BEL}, K' \rangle$, where K' is a DRS of the form exemplified in (17). ((17) is the content representation of a belief that *A* could express by means of sentence (16a).)



In general, the content representation K' of an exported belief will be a DRS with an empty Universe and a single DRS Condition, in which the ‘in-the-world-of-*T*’ operator to ‘ IN_T ’ is applied to a DRS K that is obtained by exportation from the content representation of the agent’s *T*-dedicated compartment.³⁴

³⁴This is a way of adopting, within the present framework, the idea put forward in Lewis

In the form in which it is given, the content representation of (17) fails to determine a propositional content on its own. This is because no values are provided for the drefs f' , r' , d' . When the DRS K is obtained by simple excision from the compartment-internal content representation of T , then these drefs will be the distinguished drefs of the compartment-internal ERs for the names *Frodo*, *the Ring* and *Mt.Doom* and by themselves these have no significance outside the compartment dedicated to T . But it is not hard to see how this problem can be fixed. The compartment-internal ERs can be exported and the representation of the exported belief can then make use of the distinguished drefs of those exported ERs. In other words, f' , r' , d' must be the distinguished drefs of these exported ERs, not the distinguished drefs of their compartment-internal counterparts. Thus the content representation of the belief cannot be literally the part excised from the compartment-internal content representation of T , but must be the result of replacing the distinguished drefs of compartment-internal ERs by the distinguished drefs of their exportations. The result will then be well-defined when – but also *only* when – the external belief occurs in the company of those ER duplicates.

But even when the values of the free drefs of such belief representations are defined in this way, there still remains the question what are the truth conditions of those representations. One answer that might come to mind, given how such belief representations are obtained on the basis of the agent's fiction-dedicated compartments, is that the belief is true if its content is entailed by the content of what is in her fiction-dedicated compartment. For instance, A 's belief in (17) would be true according to this answer if it is entailed by her representation of *The Lord of the Rings*. Right about this suggestion is that the values of the free drefs in the representation of the exported belief are connected in the intuitively correct way with the compartment-internal ERs that are part of the interpretation that A has constructed when she read the book: the internal ERs are members of the ER networks that are assigned to these drefs as values. Note by the way that here the way in which the fictional characters represented by ERs are defined matters. To make this fully clear it would be necessary to develop the truth conditional semantics for sentences like (16a) formally. But for that we would need to introduce much further machinery, which is out of the question here.

(1978) that 'is true in the story' can be represented with the help of a modal operator. I will briefly return to this below.

So the best I can do is to give a few hints, in the footnote attached here: ³⁵.

Even if the suggestion of the last paragraph is right on the point of dref coordination, it cannot be what we want. What it says is that an exported belief in A 's mental state is true iff it is true according to the agent's own representation of fictional content in the compartment dedicated to that content. It would be extremely difficult for such a belief ever to be false in this sense. What we want instead as truth conditions for such beliefs is that they are true according to what fictional text in question 'really says'. But is there such a thing as 'what a fictional text really says'? There are those who seriously doubt this and who may even scoff at anyone who raises the question as betraying hopeless naiveté. But it is a notion, I contend, that we cannot get around. If our agent A has the belief that Frodo threw himself into Mount Doom with the ring on his finger, then that belief is clearly false. Anyone who has read the story and has it freshly before their mind will confirm this.

I will therefore assume that for any given fictional text T there is a *basic canonical representation* $K_{T,can}$. This representation is like the representations for non-fictional texts proposed in Section 3.6 in that it has pseudo-ERs associated with each of its name tokens. But now this will be true for the fictional names of T as well as its non-fictional names (in case there are any of those).

The truth conditions of an exported belief like that in (17) can now be

³⁵ Among the things that a formal semantics for (17) must provide is a definition of the entailment relation between the compartment-internal content representation of T and the content representation in (17). A formal definition of this relation must in some way coordinate the distinguished drefs of the exported ERs that occur in the belief representation with the distinguished drefs of the corresponding internal ERs in the content representation of T . This coordination can be defined via the semantic values of the exported ERs. According to def. (15) these values are the fictional characters $[[ER']]_{N,t}$, where N is the label of the exported ER ER' . The fictional character $[[ER']]_{N,t}$ represented at t by ER' contains the corresponding compartment-internal representation ER , whose distinguished dref is the dref in the content representation of T .

For the case of (17), where the content representation of the belief is obtained by exportation from the content representation of T , this use of the ER networks that we have adopted as fictional characters may seem unnecessary, as the needed coordination between drefs in the content representation of the belief and drefs in the content representation of T is fixed by the exportation process itself. But this is no longer so for fiction-related beliefs that are not the result of exportation. Here the work done by the networks represented by the exported ERs cannot be done in any other way. For the generalization to such beliefs see below. The same is true for the further generalization to an account of the truth conditions of sentences like (16a).

stated in relation to the canonical representations of the corresponding text. For the case of (17):

- (18) (17) is true if the content represented by its DRS is entailed by the content represented by $K_{L.o.t.R,can}$, and false if the content of its DRS is refuted by $K_{LotR,can}$.

As stated, (18) conceals the role played by the operator IN_{LotR} ³⁶. But informally the role of IN_{LotR} can be described as follows. This role is two-fold. First, IN_{LotR} relates the evaluation of the DRS K in its scope to $K_{LotR,can}$, in the sense of entailment and refutation. But as part of this IN_{LotR} has a second function, that of binding the distinguished drefs occurring in K to those of the corresponding pseudo-ERs of $K_{LotR,can}$ (where 'corresponding means: belonging to the same ER network; see the remarks in footnote 35). It is this binding function of the IN-operators that transforms the representations in their scope into the forms that are required by the relations of entailment and refutation spoken of in (18).

Note that definition (18) is not only applicable to beliefs that have been formed through exportation of content DRS and ERs from a T -dedicated compartment. It can be applied irrespective of how the DRS K in the scope of the 'in-the-story' operator IN_T has been constructed, so long as the free drefs of K are distinguished drefs of exported duplicates of ERs from the T -dedicated compartment. For any such belief of in-the-story content (18) provides 'objective' truth conditions'.

(18) can also be extended to cover sentences like (16a), taken as in-the-story statements rather than as beliefs of individual story readers. There is more than one way to formulate such an extension, but an intuitive one is this: Assume a kind of ideal member A_{can} of the language community who constructs, using the principles of MSDRT, an ADF representation for the given statement in which each name N is represented by an N -labeled ER. This ER is assumed to be part of the network of N -labeled ERs and thus linked to the pseudo-ERs for the occurrences of N in the canonical representation of the story text T . (The link can be either direct, when the construction is triggered by A_{can} 's reading the statement in T , or indirectly, when A_{can} is interpreting an in-the-story statement made by someone else.). Once again the operator IN_T will then connect the distinguished drefs of the

³⁶(18) is a shorthand for a more complex statement of the truth and falsity conditions (17), which will not be given here for lack of formal machinery.

ERs in A_{can} 's representation to the distinguished drefs of the pseudo-ERs from the canonical representation $K_{T,can}$.

4.5.1 Mixing bits of in-the-story statements and meta-fiction

The point of allowing for representations of beliefs about what is true according to a text T outside the T -dedicated compartment is to provide a level of representation where such a beliefs can combine and interact with meta-fictional beliefs. Such combinations and interactions are common enough, both in our thoughts and in our fiction-related utterances. And they present what is perhaps the greatest challenge for a semantics of fiction. The statements (16c) and (16d), repeated below, illustrate some of the many phenomena that such a semantics will have to deal with.

- (16) c. You said last night that Frodo was taller than Sam.
- d. At that point Gollum bites off Frodo's finger. I think that was too much of a *Deus ex Machina*. But I loved the end nonetheless.

(16c)

gives rise to two distinct questions that have to do with truth: (i) Is that which (16c) reports you to have said true? (ii) Is it true that you said this last night? The first question allows for three possible answers – (a) yes, when the content of the complement clause of *say* is entailed by $K_{LotR,can}$, (b) no, when $K_{LotR,can}$ refutes the complement clause and (c) indeterminate, when the clause is neither entailed nor refuted. The second question allows for only two answers: (a) yes when you did say this last night and (b) no when you didn't. The first question can be settled along the lines of (18). The main problem connected with the second question is to find a suitable characterization of the direct object of the saying (i.e. the content of the complement clause). Assuming that you did say something last night about Frodo, for what words you could have used would the report (16c) be true and for what words would it be false?

Example (16d) poses problems of a different sort. Its logical form is a conjunction of an in-the-story conjunct and two metafictional ones. According to our assumptions the first is subject to a partial truth value regime and the next two to a bivalent one. What is the logic of this and other combinations of clauses that are subject to these distinct regimes? For conjunctions like (16d) this may seem a manageable problem, given all the work on partial-

valued and many-valued propositional logics that has been done within mathematical and philosophical logic. But in-the-story and meta-fictional elements can be interwoven in many other ways, for which the existing work on non-bivalent logics will be of no direct help.

As a matter of fact the difficulties that (16d) presents start even earlier, at the level of the syntax-semantics interface. How do we construct a semantic representation for this discourse from a syntactic parse of it? Salient problems are the representation of the contribution of the anaphoric pronoun *that* in the second sentence and the phrase *the end* in the third. I cannot go into this here – space is up! – but I would like to invoke the help of all who are willing to put their minds to the analysis of this example and others that they have found in fiction-related discussions or have thought of themselves.

In conclusion to Section 4.5 let me emphasize something that will have become evident in any case. The explorations of this section are just a very first timid step in the direction of a proper truth-conditional semantics for fictional and fiction-related discourse. All I have done is to look at a few examples, which were chosen more or less at random. Our look at those examples made it possible to identify some of the problems that such a semantics for fiction-related language and thought has to deal with. But there is no reason to think that we have identified more than a handful from a much larger range of issues. In particular, we should be prepared that the various ways in which in-the-story parts and meta-fictional parts can be mixed together will confront us with logical and semantic problems that are very different from anything that is known to us from semantics or formal logic. It is here that we may expect the greatest challenges to logic and semantics as we know them. Here the hard work has hardly begun.

Sobering as these thoughts may be, it should be kept in mind that accounting for the truth conditions of fiction-related language and thought hasn't been a central goal of this chapter. The reason for our explorations of truth-conditional issues in Section 4.5 was that we wanted to see how the central proposal of the chapter – to define fictional characters in terms of ER networks and to make these the denotations of fictional names – fares as part of a compositional account of truth-conditions. On this point our explorations have been instructive. For meta-fictional talk of the kind exemplified by (16b) our proposal is, we saw, of no special benefit; here almost any way of defining the denotations of fictional names will do as well as any other. But in our account of the truth conditions of in-the-story statements our proposal proves helpful. Our fictional characters – networks of ERs labeled by

fictional names – provide a simple and natural way of linking drefs representing fictional names in the semantic representations of such statements with those drefs in canonical text representations to which they must be linked in order to secure the desired truth conditions. Moreover, we may expect that assuming ER networks as denotations for fictional names will prove helpful in a compositional semantics for mixed cases like (16c) and (16d).

5 Conclusion

I limit myself to a brief summary. The principal aims of the chapters are to provide a semantics for names in fiction and for discourse containing such names. Our guiding thought has been that our use of fictional names has much in common with our use of non-fictional names, especially names of people and other entities that belong to a distant past. I have argued that what is common between fictional and non-fictional names can be captured in terms networks of Entity Representations belonging to different members of the language community. The central proposal of the chapter is to adopt the ER network associated with a fictional name N as its denotation and to identify it with the fictional character that N can be used to talk about. This, I contend, is a good way of bringing out how much our use of fictional names has in common with our use of non-fictional names, without ignoring the fundamental differences.

In the final part of the chapter (Section 4.5) some steps are taken towards a truth-conditional semantics for discourse with fictional names. The primary purpose of this investigation was to evaluate what use there might be in such a semantics for our definition of fictional characters as ER networks. The evidence we found seems modest, but it remains to be seen whether this is due to my particular choice of examples. In any case, I believe that what the chapter has to say about a truth-conditional semantics for fiction-related discourse can serve as a starting point for a more systematic approach.

My main concern in writing this chapter for the present volume has been to try to help with bridging the gap between literary studies and (formal) semantics. But, to repeat something I said early on in the chapter, the representations it proposes of what readers do with fictional texts are from a literary studies perspective quite simple-minded; they lack nearly all the structural articulation that the literary studies community will want. Of the many tasks that this essay has left incomplete or done no more than point

at, filling this hiatus is in my own view the most important.

References

- Franke, M. and de Jager, T. (2011), Now that you mention it: Awareness dynamics in discourse and decision, *in* A. Benz, C. Ebert, G. Jäger and R. van Rooij, eds, ‘Language, Games, and Evolution’, Springer, Heidelberg, pp. 60–91.
- Heim, I. (1982,1988), *The Semantics of Definite and Indefinite Noun Phrases*, Garland Press.
- Kamp, H. (1990), Prolegomena to a structural account of belief and other attitudes, *in* C. A. Anderson and J. Owens, eds, ‘Propositional Attitudes—The Role of Content in Logic, Language, and Mind’, University of Chicago Press and CSLI, Stanford, chapter 2, pp. 27–90. Reprinted in: K. von Heusinger and A. Ter Meulen, ‘Meaning and the Dynamics of Interpretation. Selected Papers of Hans Kamp’. Brill, 2013.
- Kamp, H. (2003), Einstellungszustände und einstellungszuschreibungen in der diskursrepräsentationstheorie, *in* U. Haas-Spohn, ed., ‘Intentionalität zwischen Subjektivität und Weltbezug’, mentis, Paderborn, pp. 209–289.
- Kamp, H. (2015), ‘Using proper names as intermediaries between labelled entity representations’, *Erkenntnis* **80**, 263–312.
- Kamp, H. and Reyle, U. (1993), *From Discourse to Logic*, Kluwer.
- Kamp, H., van Genabith, J. and Reyle, U. (2011), Discourse representation theory: An updated survey, *in* D. Gabbay, ed., ‘Handbook of Philosophical Logic, Vol. XV’, Elsevier, pp. 125 – 394.
- Kripke, S. (1980), *Naming and Necessity*, Harvard University Press, Cambridge. First appeared in: D. Davidson and G. Harman(eds), *Semantics of Natural Language*, Reidel, 1972.
- Lewis, D. (1978), ‘Truth in fiction’, *American Philosophical Quarterly* pp. 37–46.

- Maier, E. (2015), ‘Referential dependencies between conflicting attitudes’, *Journal of Philosophical Logic* **46**, 141–167.
- Maier, E. (2016), ‘Attitudes and mental files in discourse representation theory’, *Review of Philosophy and Psychology* **7(2)**, 473–490.
- Maier, E. (2017), ‘Fictional names in psychologistic semantics’, *Theoretical Linguistics* **43**, 1–45.
- Recanati, F. (2012), *Mental Files*, Oxford.
- Walton, K., ed. (1990), *Mimesis as Make-Believe: On the Foundations of the Representational Arts*, Harvard University Press, Cambridge, Mass.
- Wiggins, D. (1967), *Identity and Spatio-Temporal Continuity*, Blackwell.
- Wiggins, D. (2001), *Sameness and Substance Renewed*, Cambridge University Press.
- Young, J. (2018), Temporal and modal logic, in ‘Stanford Encyclopedia of Philosophy’, Stanford University.