



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di

FILOSOFIA, SOCIOLOGIA, PEDAGOGIA E PSICOLOGIA APPLICATA - FISPPA

CORSO DI DOTTORATO DI RICERCA IN: Filosofia

CICLO XXIX

The Notion of Truth through Dialetheism, Deflationism and Fictionalism

Coordinatore: Ch.ma Prof. Francesca Menegoni

Supervisore: Ch.mo Prof. Pierdaniele Giaretta

Valutatori: Ch.mi Proff. Francesco Berto e Concha Martínez-Vidal

Dottorando: Valentina Braido

Contents

Acknowledgments	i
Abstract	iii
Introduction	1
1. Dialetheism	5
1.1 Introduction to Dialetheism	5
1.1.1 Paraconsistent Logics	5
1.1.2 The Dialethic Theory of Truth	7
1.1.3 Logic of Paradox	9
1.2 Other Methodological Requirements	11
1.2.1 Disjunctive Syllogism and <i>Modus Ponens</i>	11
1.2.2 Quasi-valid Rules	14
1.2.3 The Principle of Non-Contradiction	15
1.3 Acceptance and Rejection	18
1.3.1 The Traditional View	18
1.3.2 The Dialethic View	18
1.4 The Liar	19
1.4.1 Logical Paradoxes	19
1.4.2 The Liar Paradox	20
1.4.3 Classical Theories vs Non-Classical Theories	21
1.4.4 The Dialethic Solution	24
1.5 Problematic Aspects	26
1.5.1 Revenge	26
1.5.2 The Expressibility Problem	29
2. Deflationism about Truth	33
2.1 Introduction to Deflationism	33
2.1.1 The Deflationary Theory of Truth in a Nutshell	33
2.1.2 The Rule Thesis: The Equivalence Principle	37
2.1.3 The Property Thesis: Truth Is Not a Substantive Property	40
2.1.4 The Term Thesis: Truth as a Disquotational Device	42
2.1.5 The Concept Thesis: Conceptual Connections	45
2.2 Varieties of Deflationism	46
2.2.1 Disquotationalism	46
2.2.2 Minimalism	49
2.2.3 Prosententialism	51
2.2.4 Other Varieties of Deflationism	53
2.3 Problems	54
2.3.1 Deflationism and The Generalization Problem	55
2.3.2 Deflationism and Truth-Value Gaps	57
2.3.3 Deflationism and the Liar	60

3. Truth-Theoretic Fictionalism	67
3.1 Fictionalist Theories	67
3.1.1 Introduction to Fictionalism	67
3.1.2 First Fictionalists	69
3.1.3 Good Without Being True	72
3.1.4 Fictionalist Attitude	73
3.1.5 What Fictionalism Is Not	75
3.1.6 Truth <i>In</i> and <i>About</i> the Fiction	76
3.2 Varieties of Fictionalism	78
3.2.1 Hermeneutic vs Revolutionary Fictionalism	78
3.2.2 Comparative vs Philosophical Fictionalism	80
3.2.3 Prefix-Fictionalism vs Pretense-Fictionalism	82
3.2.4 Yablo's Distinction	85
3.3 Motivations for Fictionalism	90
3.3.1 Ontological Economy	91
3.3.2 The Paradox of Existence	92
3.3.3 The Oracle	93
3.4 Truth-Theoretic Fictionalism	95
3.4.1 Fictionalism about Mathematical Discourse	96
3.4.2 Fictionalism about Truth-Theoretic Discourse	98
3.4.2 More on Alethic Fictionalism	101
4. Do Dialetheism and Deflationism Need Each Other?	105
4.1 Introducing the Issue	105
4.1.1 Can Deflationists Be Dialetheists?	106
4.1.2 Should Deflationists Be Dialetheists?	108
4.2 Should Dialetheists Be Deflationists?	110
4.2.1 Priest vs Beall: Dialethic Theories compared	111
4.2.2 Gluts and Gaps	115
4.2.3 The Number of Gluts	117
4.2.4 Contradictions Should Not Be Multiplied Beyond Necessity	120
4.2.5 From Semantics to Ontology	124
4.3 Should Deflationists Really Be Dialetheists?	125
4.3.1 Field's Paracomplete Theory	125
4.3.2 Beall vs Field: Deflationary Theories Compared	128
4.3.3 Curry's Paradox	134
4.3.4 Acceptance and Rejection	137
4.3.5 The Revenge Phenomenon	141
4.3.6 The Expressibility Problem	148
4.3.7 Conclusion	151
5. Alethic Fictionalism, Aboutness and The Liar	155
5.1 Constructive Methodological Deflationism	155
5.1.1 Introductory Remarks	155
5.1.2 Beall's Account	156
5.2 Should Dialetheists Be Fictionalists about Truth-Talk?	160
5.2.1 The Inconsistency Theories of Truth	161
5.2.2 Inconsistency Theories and Alethic Fictionalism	166

5.2.3	Dialetheism and Alethic Fictionalism	169
5.2.4	Conclusion	171
5.3	Should Alethic Fictionalists Really Be Dialetheists?	172
5.3.1	Background	172
5.3.2	Alethic Fictionalism is an Agnostic View	174
5.3.3	Conclusion	178
5.3	A New Perspective About The Liar	180
5.4.1	Introducing Yablo's Account	180
5.4.2	Semantic Excuses	180
5.4.3	Partial Truth	184
5.4.4	Subject Matter	186
5.4.5	Aboutness and Sentential Subject Matter	188
5.4.6	The Part of A about m	192
5.4.7	Yablo's Theory in a Nutshell	193
5.4.8	Tying Aboutness and Fictionalism	195
5.4.9	What <i>about</i> the Liar Paradox?	199
5.4.10	Conclusion	205
	Conclusion	207
	Bibliography	213

Acknowledgements

I wish to thank everyone who helped me write my dissertation. It was a pleasure to have Prof. Pierdaniele Giaretta as a supervisor, to whom I'm deeply grateful for having always supported and followed me in the writing of the dissertation. Thanks to University of Padova and FISPPA department for having supported my research. Thanks to University of Amsterdam and ILLC department for having hosted and given me the opportunity to attend all their courses, seminars and workshops. Thanks to my referees, Profs. Francesco Berto and Concha Martínez-Vidal for having entirely read and evaluated the dissertation. I'm very grateful to Prof. Francesco Berto also for having supervised me among ILLC and for his valuable advices both on the structure and the content of this work. I'm very thankful to Matteo Plebani for having introduced me in the magic world of Aboutness and for the countless talks on almost all the topics of my dissertation. Thanks to everyone who contributed to the success of this work with suggestions and comments over coffee and spritz. In particular, thanks to Riccardo Baratella, Prof. Massimiliano Carrara, Roberto Ciuni, Francesco Gallina, Graham Priest and Beppe Spolaore. Finally, I'd like to thank the library of philosophy of the University of Padova for the professionalism and kindness of its staff.

Abstract

English

In the present work three different areas about the truth predicate coexist: the *semantic* area, whose focus is on theories dealing with the solution to the semantic paradoxes that involve the truth predicate; the *ontological* area, which covers the theories investigating the nature of truth; and the area of *discourse*, in which the focus is on those theories that analyse how the discourse about the truth predicate has to be understood. The aim of this work is to analyze the consequences arising from certain specific ways in which these areas can interact with each other. The first three chapters expose the main features of the theories corresponding to the three areas about the truth predicate, respectively dialetheism, deflationism and fictionalism. The possible combinations between those theories are the focus of the two following chapters. Finally, a new attempt to account for the liar paradox is provided.

Italiano

Il presente lavoro interseca tre diversi ambiti concernenti il predicato di verità: l'ambito *semantico*, che vede al centro le teorie che si occupano della risoluzione dei paradossi semantici che coinvolgono il predicato di verità; quello *ontologico*, nel quale sono interessate le teorie che si interrogano su quale sia la natura della verità; e l'ambito, per così dire, *del discorso*, all'interno del quale si muovono le teorie che analizzano in che modo debba essere inteso il discorso riguardo al predicato di verità. Lo scopo di questo lavoro è analizzare le conseguenze che emergono da alcuni specifici modi in cui questi tre ambiti possono interagire tra di loro. Nei primi tre capitoli sono espone le caratteristiche peculiari delle teorie corrispondenti alle tre aree del predicato di verità, rispettivamente dialeteismo, deflazionismo e finzionalismo. Le possibili combinazioni tra queste teorie saranno l'argomento dei due capitoli successivi. Infine, il lavoro si conclude con un nuovo tentativo di rendere conto del paradosso del mentitore.

Introduction

In the present work three different areas about the truth predicate coexist: the *semantic* area, whose focus is on theories dealing with the solution to the semantic paradoxes that involve the truth predicate; the *ontological* area, which covers the theories investigating the nature of truth; and the area of *discourse*, in which the focus is on those theories that analyse how the discourse about the truth predicate has to be understood. My purpose here is to analyze the consequences arising from certain specific ways in which these areas can interact with each other.

For each of the previous areas about the truth predicate I'm going to take into account a correspondent theory and to analyze how these three theories can be combined and, in particular, I wish to figure out whether there are effective advantages in their union, or not. The semantic area will be represented by dialetheism, the ontological area will affect deflationary theories of truth and, lastly, the area of discourse about truth will be focused on fictionalism about truth-talk.

Point of contact for the three theories is the well-known Liar paradox, that is, a sentence that says of itself that is false. The liar can be formalized as follows:

(L) (L) is false.

Let's reason by cases. If we assume (L) to be true, then what it says is true, but it says that it is false, hence, if (L) is true then is false. Hence, implicit contradiction. If we assume, now, (L) to be false, then this is exactly what (L) says and, hence, it turns out to be true. So, if (L) is false then is true. Implicit contradiction again. Both alternatives lead to contradiction. Hence, (L) is a paradox. We will see that every theory of truth will face the liar sooner or later, be it a theory from the semantic area, the ontological or the area of discourse.

The first chapter is focused on *dialetheism*, the theory developed by Graham Priest, according to which there are true contradictions, called *dialetheias*. A dialetheia, thus, is a sentence A such that both A and its negation are true. Another way to define a dialetheia is as a sentence that is both true and false at the same time. The main problem for one who accepts dialetheias is *trivialism*.¹ In order to avoid trivialism, dialetheism subscribes a paraconsistent logic, namely, a logic in which the rule known as *ex contradictione quodlibet* doesn't hold. This rule, known also as *explosion*, maintains that from a contradiction everything follows: for every A and B, $A, \neg A \vdash B$. As a result, according to dialetheism not every contradiction is true but only few of them are, contrariwise to trivialism. In the first part of the chapter I'm going to outline the theory developed by Priest in details, by putting the spotlight on the solution dialetheists provide to semantic paradoxes and, in particular, to the liar.

¹ A theory is *trivial* if every sentence of the theory is true.

In the second part of the chapter, I'll focus on the two main problems affecting a dialetheic theory, i.e. the revenge phenomenon of the liar and the problem of the impossibility for a dialetheist to express certain fundamental notions of his theory within his own language.

The topic of the second chapter is *deflationary* theories of truth that belong to what I called the ontological area about the truth predicate.² These theories arise from the attempt of providing an answer to the question "What's the nature of truth?" According to deflationism, the truth predicate expresses no substantive property of truth at all. In other words, there's nothing about the nature of truth that goes beyond what is captured when we say, for instance, that "the snow is white" is true if and only if snow is white. To say of a sentence that is true means nothing but asserting the sentence itself and, hence, the truth predicate adds nothing to the meaning of a sentence. In this chapter I'm going to explain the main theses a deflationary theory endorses: the rule thesis, according to which the use of a truth-bearer and the use of an ascription of truth to it are intersubstitutable in all transparent contexts; the property thesis which holds that there is no genuine substantive property of truth; the term thesis that emphasises the merely expressive role of the truth predicate; and the concept thesis, which holds that the concept of truth cannot be clarified at a more fundamental and deeper level via the use of other concepts which the truth predicate is connected to. After having emphasised the reasons for assuming the deflationary theories of truth and the possible divisions we can make inside them, the chapter will end with the attempt to highlight the main problems and weaknesses of deflationism, with particular attention for those about the liar.

The third chapter is focused on *fictionalism* and, in particular, on fictionalism about truth-talk. Fictionalism about a region of discourse is the view that the assertions made inside that region are not literally true, but rather they are only fictionalistically true. In other words, the sentences belonging to that area of discourse must be taken as similar to sentences made into a fictional book. If we consider, for instance, mathematical fictionalism, the idea behind it is that sentences such as " $2+2=4$ " have some very evident analogy with sentences like "Sherlock Holmes lives in London, in Baker Street 221b." According to mathematical fictionalism both sentences are false if taken at face value, because there is no object which the name "Sherlock Holmes" refers to and, in the same way, there is no object which names like "2" and "4" refer to as well. However, roughly speaking, the two sentences are fictionalistically true, that is, they are true if considered as parts of a specific fiction. The result is that the latter is true according to the stories of Sherlock Holmes and, similarly, the former is true according to the story of standard mathematics. Alethic fictionalism is the theory that the discourse about truth works through elements of fiction. In the first part of the chapter I'm going to outline the general features of fictionalist theories and in the second part I'll focus on truth-

² It's not properly correct to define deflationism as an ontological theory of the truth predicate, because deflationists maintain that the truth predicate has no underlying nature at all and, for that reason, we should move from asking "What's truth?" to asking "What does 'true' do within our language?" I maintain that deflationism deserves the label of ontological theory simply because before she analyses the functioning of "true" within the language, a deflationist, in fact, tells us what our ontological commitment towards the truth predicate must be and, hence, *in primis* deflationism can be considered as an ontological theory of the truth predicate.

theoretic fictionalism and, in particular, on the main argument in favor of such a perspective of truth-talk.

Aim of the fourth chapter is to intertwine the first two areas of the truth predicate through their correspondent theories: the ontological area represented by deflationism and the semantic area embodied by dialetheism. A deflationist should be dialetheist, this is what Bradley Armour-Garb and J.C. Beall maintain. Starting from this claim, I'll try to provide an answer to two complementary questions: "Assuming that deflationists should be dialetheists, is the opposite also true? In other words, should dialetheists be deflationists as well?" and "Should deflationists really be dialetheists, as Armour-Garb and Beall maintain?"

To answer the former question, I'm going to compare two different dialetheic theories: the theory developed by Priest, which doesn't assume a deflationary approach of the nature of truth; and the theory recently developed by Beall that is the perfect example of deflationary dialetheism. The cost-benefit analysis of the two theories will allow me to provide an answer to the former question.

The answer to the latter question, instead, will require first an answer to the following question: "One of the main reasons why Hartry Field developed his own theory of truth is to have a theory fitting very well with deflationism. Why doesn't Field rely on dialetheism in the development of a semantic for his deflationary theory and, on the contrary, prefers to develop his own and more complex paracomplete theory?" Hence, also the second part of the chapter requires a comparison between two theories: the deflationary and dialetheic theory of Beall, and the deflationary and paracomplete theory of Field. A theory is paracomplete if the law of excluded middle doesn't hold in it: for every A , $A \vee \neg A$. Also in this case, the cost-benefit analysis of the two theories is what will enable me to provide an answer to the second question of the chapter.

My point in this chapter is that both questions deserve a negative answer. This means two things: on the one hand, that dialetheists shouldn't be deflationists, not because they should be non-deflationists, but rather because the dialetheic logic and semantics fit well also with other theories about the nature of truth; and, on the other hand, that deflationists shouldn't be dialetheists, contrariwise to what asserted by Armour-Garb and Beall, not because they should be non-dialetheists, but rather because in order to solve the liar paradox they have at least two others available possibilities.

The fifth chapter brings alethic fictionalism into the debate. The aim of this chapter is twofold: in the first place, I wish to figure out whether fictionalism can get along with dialetheism, or not; and, on the other hand, I wish to focus on a new approach about meaning recently developed by Stephen Yablo. In the first part of the chapter, starting from a paper published by Beall in which the author defends a particular perspective that tries to keep all the three theories here analysed – i.e. dialetheism, deflationism and fictionalism – together, my aim is, again, to answer two different and complementary questions: "Should dialetheists be alethic fictionalists?" and "Should truth-theoretic fictionalists really be dialetheists, as Beall claims?" Both questions will have a negative answer. The answer to the former question will be obtained through the comparison between dialetheism and inconsistency theories of truth, by putting the spotlight on the reasons why the advocates of the latter

Introduction

assume a fictionalist perspective of truth-talk and by arguing that these motivations cannot be shared by dialetheists as well. In order to answer the latter question, instead, I'll emphasise that other solutions to the liar paradox, different from the dialethic one, are available to the alethic fictionalist.

In the second part of the chapter I'm going to outline the main features of the aboutness' theory developed by Stephen Yablo. Then, I'm going to achieve two fundamental goals: on the one hand, I'll try to show that Yablo's theory of aboutness can be translated into a fictionalist theory, or, in other words, that the key notions of fictionalism can be translated into the main concepts of aboutness' theory, and vice versa; and, on the other hand, I'll put forward an attempt of applying Yablo's theory to the problem of the liar paradox. This latter attempt is destined here to remain so, but, at the same time, it offers the possibility for further investigations and future developments.

Chapter 1

Dialetheism

1.1 Introduction to Dialetheism

1.1.1 Paraconsistent Logics

The word “paraconsistent” literally means “beyond the consistent” and it was coined in 1976 by the Peruvian philosopher Miró Quesada.¹ Paraconsistent logics belong to the broader category known as non-classical logics because they reject some typical laws and rules of classical logic. The main feature of paraconsistent logics is the rejection of the classical logic’s rule known with the name of *Law of Scoto*,² so much so that a logic is paraconsistent if and only if (henceforth, iff), by definition, it rejects such rule. The law of Scoto is also known with the medieval name of *Ex Contradictione Quodlibet* (ECQ), which literally means that from a contradiction everything follows, and it can be formalized in the following way:

$$(ECQ_1) \quad A, \neg A \vdash B$$

or the equivalent:

$$(ECQ_2) \quad A \wedge \neg A \vdash B$$

for every A and B . The two formulations respectively correspond to the *distributive* and *collective* formulations of a contradiction. In the former case, the contradiction consists of a couple of sentences, one of which is the negation of the other; and, in the latter case, the contradiction consists of the conjunction of two sentences, one of which is the negation of the other. In classical logic and in non-classical ones that I’ll refer to in this place, there is no difference between the two formulations.³

The relation of logical consequence is called *explosive* iff it validates (ECQ₁), and a logic is *explosive* iff its logical consequence is. This means that a logic is explosive iff the law of Scoto holds in it, which law is also known with the name of *explosion*, because, to use a metaphor due to Berto, it results in a

¹ Quesada[1976].

² It is also called *Law of Pseudo-Scoto*, because the rule has been wrongly attributed to the Scottish philosopher Duns Scoto, but it is actually due to an anonymous author.

³ Examples of logics in which this equivalence doesn’t hold are the so-called non-adjunctive systems, which reject the adjunction rule ($\{A, B\} \models A \wedge B$), namely, the rule that allows us to move from a distributive contradiction to a collective one. For a brief but effective introduction to the topic of non-adjunctive systems see Berto[2006 and 2007] and Priest, Tanaka and Weber[2015].

real conceptual explosion that makes an inconsistent logic philosophically useless.⁴ More specifically, it's worth noting that B in (ECQ₁) could be any well-formed sentence of the language, also a contradiction itself; this means that from a contradiction an infinite number of them follows. An example of how (ECQ₁) works is the following: from the sentence “pigs fly and pigs don't fly” we can legitimately (according to the rule) conclude that the Earth goes around the sun. This specific case does not give rise to any problems, but from the same sentence we can also infer (by (ECQ₁)) that the Earth doesn't go around the sun, and also that the Earth goes and doesn't go around the sun at the same time. These last two conclusions are more problematic because they are plainly false. The view that all sentences of the language (therefore also all contradictions) are true is called *Trivialism*. In other words, a system is trivial if it allows proving all the sentences of its language. Accordingly, the conclusion is that if the law of Scotto holds, then from a true contradiction trivialism immediately follows. This means that in classical logic there is no difference between inconsistency⁵ and trivialism.⁶ In general, this is not the case, because while every trivial system is also inconsistent, the opposite is not the case. An example of an inconsistent non-trivial system is the one that gives the title to this paragraph, that is, paraconsistent logics, which, by rejecting (ECQ₁), are known to accept inconsistency without being trivial. The words of the English philosopher and logician Graham Priest comes in favour of the failure of (ECQ₁):

The failure of explosion is a plausible logico-metaphysical view, and that one cannot simply assume otherwise without begging the question.
[Priest 1998, p.416]

Priest underlines that the acceptance of the explosion rule can only take place by assumption, and using this assumption against paraconsistent logics it would beg the question against them.

The methodological requirement of rejecting (ECQ₁) for a paraconsistent theory inevitably leads to other important changes in logic, the discussion of which is delayed to section 1.2. What was interesting to point out in the present section is how it can be possible to accept a true contradiction inside our own logical system without falling into trivialism and we saw that this task can be achieved by relying on a paraconsistent logic.

⁴ See Berto and Bottai[2015], p.50.

⁵ *Consistency* is the property of a formal system provided with negation, according to which we can never deduce in the system both a formula and its negation, that is, a contradiction. If, instead, this happens, then the system is called *inconsistent*, and it is said that it has the property of *inconsistency*.

⁶ Good evidence of this can be found in the name that sometimes trivialism acquires, that is, “absolute inconsistency”.

1.1.2 The Dialetheic Theory of Truth

Dialetheism is the view that there are true contradictions, that is, sentences that are true and false at the same time.⁷ In other words, this is equivalent to say that dialetheism allows for both a sentence and its negation to be true. The name *dialetheism* comes from the term *dialetheia* that was coined by Graham Priest and Richard Routley in 1982.⁸ It doesn't come from any other name already existing in any other natural language, and the reason is:

After exhausting all the dictionaries at our disposal (including Greek, Russian and Gaelic), we decided that no extant word would express this idea. So we were forced to coin one. [Priest, Routley and Norman 1989, p.XX]

and it means “two-way truth” because, to continue:

A true contradiction is a Janus-faced creature which faces both truth and falsity. [Priest, Routley and Norman 1989, p.XX]

In particular a *dialetheia* is defined as a sentence A , such that both A and its negation, $\neg A$, are true. So, quite simply, dialetheism is the view that there are *dialetheias*.⁹ As a matter of fact, the advocates of dialetheism believe that some contradictions are inevitable and demonstrable because they affect our ordinary language and our thought processes.¹⁰

It's easy to guess that, according to Priest, the inconsistency of our language is far from being a problem, and, on the contrary, is a peculiar feature we should accept as such. To use his own words:

The inconsistency of our linguistic principles is the very thesis I am affirming. Hence, in discussions of what these principles are, consistency cannot be invoked as a regulative principle without begging the question against me. But, even putting this aside, there is no legitimate presupposition of consistency here. Rather, the natural presupposition is that of inconsistency. For language and the principles that govern it have developed piecemeal and under no central direction. As logicians know, inconsistency is the natural outcome of spontaneity. Consistency has to be fought for. [Priest 1987, p.5]

According to Priest, classical logicians can support consistency only by begging the question against dialetheism and not with a legitimate argument, as it would be desirable. Let's now display the putative defective argument of classical logic

⁷ How Priest and Berto[2013] pointed out, the definition of dialetheism is given in terms of sentences, but we could formulate it in terms of any other notion we prefer as truth-bearer (i.e. statements, propositions, etc.) without making big changes.

⁸ See Priest, Routley and Norman[1989], p.XX.

⁹ For a detailed explanation of what a *dialetheia* is and which areas of our language are affected by *dialetheias* according to a dialetheist philosopher, see Priest[1987 and 1995] and Priest and Berto[2013].

¹⁰ For instance, according to dialetheism, there are also observable contradictions, like in the paradoxes of perception induced by the images and prints of the Dutch graphic artist, Maurits Cornelis Escher.

used against dialetheism.¹¹ According to classical logicians, the truth conditions for negation are the following: $\neg A$ is true iff A is not true. From now on the argument will proceed as a *reductio ad absurdum* argument. First of all, let's assume that dialetheism is true, and so let's assume the following premise: $T(A \wedge \neg A)$ for some A , where T is the truth predicate. In other words, we are assuming that some contradictions are true. By the truth tables for conjunction, this is equivalent to $T(A) \wedge T(\neg A)$, but, the classical logician says, this is not possible because of the truth conditions for negation stated above (i.e. $\neg A$ is true iff A is not true) that go against what we have obtained here, i.e. that both A and $\neg A$ are true. The conclusion of the argument is that we must reject the initial assumption, i.e. we must reject $T(A \wedge \neg A)$.

According to Priest, the problem with this argument is twofold. First of all, it is important to highlight that the classical truth conditions for negation are contentious, because it is not so obvious that they are commonly and universally accepted.¹² Secondly, Priest emphasises that the impossibility to consider both A and $\neg A$ true is a consequence of this being a contradiction. But the impossibility to accept true contradictions is exactly what the classical logician was arguing for. Hence, this argument of classical logic begs the question against dialetheism.¹³ Finally, Priest says, it is much more natural to assume that the language is inconsistent rather than consistent. After all, the semantic paradoxes seem to bring to light exactly this fact, i.e. the inconsistent character of our language. I'll come back to this point in section 1.2.3, where I'll analyse what are the consequences of the inconsistency of the language postulated by Priest, in terms of acceptance or rejection of the principle of non-contradiction.

There is a further methodological requirement for a dialetheist theory of truth that is important to specify here, namely, dialetheism is not equivalent to trivialism. Dialetheism and trivialism are not the same view: while, according to trivialism, all contradictions are true, this is not the case for dialetheism, according to which, only few contradictions, and not all of them, are true. This means that, according to dialetheism, only some sentences are dialetheias. This is precisely the reason – i.e. not to falling into trivialism and consequently not to become a useless naïve theory – why dialetheism relies on a paraconsistent logic, according to which, it might be useful to remember, a contradiction doesn't produce an explosion capable of multiplying contradictions indefinitely. Despite this fellowship between paraconsistent logics and dialetheism, it is important not to confuse them with each other. In fact, dialetheism is a theory of truth and paraconsistency is a property of the relation of logical consequence. Despite, most of the times, dialetheism and paraconsistent logic go hand in hand, we could legitimately have one without the other. In fact, for dialetheism is convenient to rely on a paraconsistent logic because of its methodological requirement of anti-triviality, but this doesn't mean that also all

¹¹ Actually, in Priest[1998b] are presented some other arguments in favor of non-contradiction, together with the respective answers given by the dialetheist philosopher. I'll consider here only one of them.

¹² In fact, as we might expect and as we will see below, Priest doesn't accept the validity of the truth-conditions for negation stated in this way.

¹³ Priest[1998b], p.418.

paraconsistent logics have to be dialetheist, and, on the contrary, this is far from the truth.¹⁴

1.1.3 Logic of Paradox

After having explained what dialetheism is from a philosophical point of view, let's now look at its logic. Priest proposes a paraconsistent logical system, called *Logic of Paradox* (LP), in which “true” and “false” are not exclusive. More specifically, LP is a three-valued semantics and the truth-values are “true”, “false” and the new one, i.e. “both true and false” (also called *paradoxical*).

In order to better understand the dialethic logic, let's consider the following quote from Priest:

Considering how radical dialetheism is, perhaps the surprising thing is how *little* its logical theory differs from the classical theory. [Priest 1987, p.110 (emphasis in original)]

So, according to Priest – and we will see he is right – the dialethic logic is not so different from classical one. What are the main differences? To use Priest's words one more time:

Essentially, it just generalizes classical logical theory to allow it to handle a domain that was, before, beyond the pale – the inconsistent. [Priest 1987, p.110]

Therefore, the only relevant difference between the logic of paradox and classical logic is that, according to the former, inconsistency is taken into account like any other hypothesis, and not simply rejected as a problem, contrariwise to the latter.

In order to see how little are the differences between classical and dialethic logic, let's begin by comparing their corresponding truth tables for the main connectives – i.e. negation, conjunction, disjunction and the conditional, which, according to both logics, is the material one.

- (1) *Negation*. In LP, like in classical logic, if A is true (false), then its negation is false (true). But, of course we have to take into account also the new truth-value: if A is paradoxical, then so is its negation.

A	$\neg A$
V	F
F	V
P	P

¹⁴ Some examples of non-dialethic paraconsistent logics are *relevance logic*, *positive-plus systems*, *non-adjunctive logics*, and so on. For a quick and brief overview of these accounts see Berto[2006 and 2007]. For other examples of non-dialethic paraconsistent logics see Priest, Tanaka and Weber[2015].

- (2) *Conjunction.* In LP, a conjunction is true if both its conjuncts are true, and false if at least one of its conjuncts is false (as well as in classical logic). Moreover a conjunction is paradoxical in all other cases.

A	B	$A \wedge B$
V	F	F
V	V	V
F	F	F
V	P	P
F	P	F
P	P	P

- (3) *Disjunction.* In LP, a disjunction is true if at least one of its disjuncts is true, and false if both its disjuncts are false (as well as in classical logic). Otherwise, a disjunction is paradoxical.

A	B	$A \vee B$
V	F	V
V	V	V
F	F	F
V	P	V
F	P	P
P	P	P

- (4) *Conditional.* Both in LP and in classical logic, the conditional is the material one, so its truth table is constructed on the basis of the tables for negation and disjunction: $A \rightarrow B =_{\text{def}} \neg A \vee B$.

A	B	$A \rightarrow B$
V	F	F
V	V	V
F	V	V
F	F	V
V	P	P
F	P	V
P	F	P
P	V	V
P	P	P

The white part of the truth tables corresponds to the behaviour of the connectives in classical logic. The entire truth tables – i.e. both the white and

the red parts – are exactly the truth tables of the dialethic logic. What we immediately note is that the truth tables of LP behave exactly in the same way as those of classical logic for the standard value “true” and “false”. Hence, the only difference lays is the new truth-value, namely, paradoxical, which solely expands the truth tables for the connectives of classical logic, and doesn’t radically change them. In this way, LP turns out to be only an extension of classical logic, because the truth tables of the latter are subsumed by those of the former. Instead of upsetting the truth tables of classical logic, LP simply allows us to effectively treat inconsistent contexts. Indeed, this is so since classical logic is just a special case of LP where no sentence takes the paradoxical value.¹⁵

This is not the only difference between LP and classical logic, but it’s the one that most clearly show the close links between them. As you can imagine, there are other important changes in logic – for instance, in LP we have to reject some central rules of classical logic, in order to avoid some unpleasant consequences –, but we will postpone this analysis to section 1.2.

Another very quick remark is worth making. In fact, we can easily note that the matrices used for the truth tables of LP are the same of Kleene logic (known as K3)¹⁶ – which is the logic used by Saul Kripke and Hartry Field in the construction of their own theories. The difference between LP and Kleene logic is the interpretation of the third truth-value: on one side, in Kleene logic the third value is called *undefined* and it corresponds to neither true nor false; on the other side, as we saw, in LP the third value is called *paradoxical* and it corresponds to both true and false. More specifically, according to Kleene, “true” is the only designated value and, on the contrary, according to LP, the designated values are both “true” and “paradoxical”.

1.2 Other Methodological Requirements

1.2.1 Disjunctive Syllogism and *Modus Ponens*

As we have already said above – in section 1.1.1 about paraconsistent logics – the methodological requirement of rejecting the explosion rule inevitably leads to other important changes in the logic of a paraconsistent theory. In particular, it leads to the rejection of other fundamental rules of classical logic that are involved exactly in the demonstration of the explosion. In order to make this clear, let’s consider the following proof for (ECQ₂):

¹⁵ The same goes if we extend LP to a quantificational logic, called LPQ. In this case we have:

- *Existential quantification.* In LP, $\exists x(A(x))$ is true if $A(x)$ is true for some assignment of an element of the domain D to x ; it is false if $A(x)$ is false for all assignments to x . (these two clauses are the same for classical logic); and paradoxical otherwise.
- *Universal quantification.* The universal quantification $\forall x(A(x))$ can be defined as $\neg\exists x(\neg A(x))$. In LP, $\forall x(A(x))$ is true if $A(x)$ is true for all assignments to x ; it is false if $A(x)$ is false for at least one assignment to x (these two clauses are the same for classical logic); and paradoxical otherwise.

¹⁶ Kleene[1952], p. 334-335.

Chapter 1

(1)	$A \wedge \neg A$	[by assumption]
(2)	A	[by 1 and \wedge -Elimination]
(3)	$A \vee B$	[by 2 and \vee -Introduction]
(4)	$\neg A$	[by 1 and \wedge -Elimination]
(5)	B	[by 3, 4 and Disjunctive Syllogism]

In order to reject (ECQ₂) – that, remember, is equivalent to (ECQ₁) – we must reject at least one of the rules involved in its demonstration: namely, the Elimination of the Conjunction (\wedge -Elimination), the Introduction of the Disjunction (\vee -Introduction) and the Disjunctive Syllogism. The first two rules regulate, respectively the meaning of conjunction and disjunction. The first one tells us that given a true conjunction, every conjunct is true as well (for instance, given the truth of “dogs bark and cats meow”, we can correctly individually infer the truth of “dogs bark” and “cats meow”). The second rule tells us that if a sentence is true, then the disjunction that has this sentence as a disjunct is true as well (for instance, given the truth of “dogs barks”, we can correctly infer the truth of a disjunction involving it, such as “dogs barks or cats barks”). Both these two rules are universally accepted and, thus, there doesn’t seem to be arguments for their rejection. Hence, the rule strongly suspected of being the cause of the explosion is the only one left, namely, the disjunctive syllogism.

Obviously, a paraconsistent logician can’t reject a rule only because it is engaged in the proof of an unwelcome rule; otherwise she risks accusation of *ad hocness*. Then, which is the freestanding reason that allows a paraconsistentist to reject the rule of disjunctive syllogism? According to Priest, it’s the fact that it doesn’t preserve truth. In fact, if we assume that both A and $\neg A$ are true, then $A \vee B$ is true as well, even if, let’s suppose, B is false, because we are using the introduction of the disjunction on A . But this means that from true premises (A , $\neg A$, and $A \vee B$) we are inferring, by disjunctive syllogism, a false conclusion (B). Therefore, as we have already said above, the disjunctive syllogism has to be rejected because it doesn’t preserve truth.

Unfortunately, there are other problems a paraconsistent logician has to face: the consequences of the rejection of the disjunctive syllogism and (ECQ₂) go beyond what we have just said and go in undesirable directions. In classical logic, the conditional is the material one, that, by definition, doesn’t take semantic, causal or intentional implications into account. As a matter of fact, the material conditional is defined by referring only to negation and disjunction:

$$(CM) \quad A \rightarrow B =_{df} \neg A \vee B$$

Hence, the material conditional tells us that either it is not the case that the antecedent, or it is the case that the consequent. Therefore $A \rightarrow B$ is false only in one situation, when A is true and B is false. The problem is that if the conditional is the material one, then we will have the equivalence between the disjunctive syllogism and another fundamental rule of classical logic, i.e. *modus*

ponendo ponens (most widely known simply as *modus ponens*).¹⁷ The latter can be formalized as follows:

$$(MP) \quad A \rightarrow B, A \vdash B$$

By the validity of the principle of double negation,¹⁸ (MP) is equivalent to the disjunctive syllogism:

$$(DS_1) \quad \neg A \vee B, A \vdash B$$

As we have just shown, if we reject the rule of disjunctive syllogism in a logic in which the conditional is the material one, then we are also forced to reject the very intuitive rule of *modus ponens*. But this rule seems to be an unavoidable requirement for any logical system because it gives exactly the intuitive meaning of the conditional, and, so, it seems not to be rejectable. On this matter, Priest himself asserts that:

As I have stressed, *modus ponens* is a *sine qua non* of any implication connective. [Priest 1987, p.86 (emphasis in original)]

And he is not the only one to assume *modus ponens* as a *sine qua non* of the ordinary reasoning. Also the American philosopher and logician Hartry Field shares the same view:

The absence of a conditional with *modus ponens* is enough to prevent any semblance of ordinary reasoning. [Field 2008, p.369 (emphasis in original)]

Which possibilities do we have now? We cannot have it both ways: either we give up the idea of a paraconsistent logic – but it doesn't seem a reasonable option, especially for a supporter of dialetheism, such as Priest is –,¹⁹ or we must give an independent semantics for the conditional independently from both disjunction and negation. In the words of Priest:

Any conditional worth its salt, \rightarrow , should satisfy the *modus ponens* principle: $\{\alpha, \alpha \rightarrow \beta\} \models \beta$. This is, indeed, analitically part of what implication is. Yet this principle fails for material implication as we saw. $\{\alpha, \neg\alpha \vee \beta\} \models \beta$ is not, in general, true. Hence, material implication is not the conditional. [Priest 1987, p.83 (emphasis in original)]

¹⁷ Actually, if the conditional is the material one, the disjunctive syllogism and *modus ponens* are equivalent to one more rule, that is, *modus tollendo tollens* (usually known as *modus tollens* for short), that can be formalized in the following way:

$$(MT) \quad A \rightarrow B, \neg B \vdash \neg A$$

¹⁸ The Principle of Double Negation states that a proposition is equivalent to the falsehood of its negation ($\vdash p \equiv \neg\neg p$).

¹⁹ As we saw, if we gave up paraconsistency, dialetheism would collapse into trivialism.

1.2.2 Quasi-valid Rules

Actually, Priest does not properly reject the rules we mentioned above – i.e. disjunctive syllogism, *modus ponens* and *modus tollens* –, but, instead, he simply restricts their use. More specifically, he proposes to consider them as “quasi-valid” rules of inference,²⁰ where a quasi-valid rule is a rule that is classically valid but dialetheically invalid. Hence, we are allowed to use a quasi-valid rule in all and only non-paradoxical contexts. So, according to dialetheism and contrariwise to classical logic, quasi-valid rules, such as disjunctive syllogism, are not generally and universally valid, because we can use them only if we are sure we are dealing with a consistent situation. This is the cost of dialetheism and it is not as great as one might expect, according to Priest.²¹ He goes on by specifying when we are allowed to use a quasi-valid rule:

A quasi-valid inference is usable in a consistent situation in the following sense: if the premises are rationally acceptable then so is the conclusion, provided the crucial contradiction is rationally rejectable. [Priest 1987, p.115]

So, when to use a quasi-valid rule is a pragmatic issue and not a formal one. We can use a quasi-valid inference if we are sure we are not dealing with a paradoxical sentence. But, of course, paradoxical sentences are not marked as such, and so, the question is: how can we recognize them? According to Priest, the dialetheias are a small portion of the contradictions and, consequently, of the sentences in general, contrary to non-true contradictions that are the majority. This low probability of dialetheias is well explained by the dialetheist:

The statistical frequency of dialetheias in normal discourse is low. Dialetheias appear to occur in a quite limited number of domains: certain logico-mathematical contexts, certain legal and dialectical contexts, and maybe a few others. Moreover, even in the domains where they do occur, very few contradictions are dialetheias. [Priest 1987, p.116]

This explains why we commonly use quasi-valid rules, such as disjunctive syllogism and *modus ponens*, and we do this successfully. If dialetheias were common, we would expect quasi-valid rules to go wrong quite frequently. Instead, the normal success of quasi-valid inferences is sufficient to show the infrequency of dialetheias.

These considerations help Priest to state a methodological maxim that explains when we are allowed to use a quasi-valid inference:

Methodological Maxim (M). Unless we have specific ground for believing that the crucial contradictions in a piece of quasi-valid reasoning are dialetheias, we may accept the reasoning. [Priest 1987, p.116]

²⁰ See Priest[1979], p.231 and Priest[1987], p.110.

²¹ Priest[1987], p.110.

In other words, it is usually reasonable to use quasi-valid rules. In fact, it seems plausible to think that in our everyday reasoning we are not dealing with dialetheias.

To conclude, according to Priest, all contexts are consistent by default. In these contexts, classical logic and dialetheism behave and reason in the same way. The difference between them is in contexts with dialetheias, where classical logic collapses into trivialism because of explosion. On the contrary, a dialetheist has the choice either of dropping the theory if he thinks that the contradiction is not acceptable; or, if he thinks it is acceptable, to keep it valid by not using any quasi-valid inference.

1.2.3 The Principle of Non-Contradiction

The Principle of Non-Contradiction (PNC) certainly is one of the parties in presence when we talk about paraconsistent logics and dialetheism. In fact, both theories admits the existence of true contradictions and so, to some extent, they seem to challenge the *principium firmissimum*, as Aristotle used to define the principle of non-contradiction.²² Let's see now if and in what respect this is correct.

As it is well know, the origin of the principle of non-contradiction goes back to the Greek philosopher Aristotle. In book lambda of *Metaphysics* and in its works about logic, Aristotle introduces two *sine qua non* principles that complement each other: the Principle of Excluded Middle (EM) and, of course, the Principle of Non-Contradiction. The former tells us that either it is the case of a sentence, or it is the case of its negation, and no third way is admitted.²³ It can be formalized as follows:

$$(EM) \quad \alpha \vee \neg\alpha$$

for every α in the language. According to the Polish philosopher Jan Łukasiewicz,²⁴ there is more than one formulation of the principle of non-contradiction that can be tracked down from Aristotle's works.²⁵ Take a look now at the different formulations of (PNC) we can identify, according to Łukasiewicz:

- The *syntactic* formulation tells us that it is not possible to have both a sentence and its negation at the same time:

²² Aristotle[1928].

²³ In classical logic, the principle of excluded middle is equivalent to the Principle of Bivalence (BV), according to which a sentence has only one truth-value, either true or false. In formal language, this is:

$$(BV) \quad T(\alpha) \vee F(\alpha)$$

This equivalence between (EM) and (BV) doesn't hold in LP. In the logic of paradox (EM) holds, but (BV) doesn't, because we no longer have a two-valued semantics and a third way is possible.

²⁴ Łukasiewicz[1951].

²⁵ See Berto[2006] and Berto and Bottai[2015].

$$(PNC_1) \quad \neg(\alpha \wedge \neg\alpha)$$

- The *logico-semantic* formulation assures us that a sentence cannot be both true and false at the same time:

$$(PNC_{2a}) \quad \neg(T(\alpha) \wedge F(\alpha))$$

In classical logic – as well as in the dialethic logic LP –, falsity is defined as the truth of negation. For this reason (PNC_{2a}) is equivalent to:

$$(PNC_{2b}) \quad \neg(T(\alpha) \wedge T(\neg\alpha))$$

This, in its turn, in classical logic – but not in LP – is equivalent to:

$$(PNC_{2c}) \quad \neg(T(\alpha) \wedge \neg T(\alpha))^{26}$$

- The *ontological* formulation prevents the existence of an object that both has and not has the same property at the same time:

$$(PNC_3) \quad \forall x \forall P [\neg(P(x) \wedge \neg P(x))]$$

- The *psychological-pragmatic* formulation forbids a sentence to be both accepted and rejected by the same person at the same time:

$$(PNC_{4a}) \quad \neg(\vdash_x \alpha \wedge \dashv_x \alpha)$$

where \vdash_x and \dashv_x are respectively the operators for acceptance and rejection, and they tell us that the person x accepts something, in the former case, and rejects something, in the latter. In classical logic – contrary to LP and other non-classical logics – the rejection of a sentence is equivalent to the acceptance of its negation.²⁷ Hence, (PNC_{4a}) is equivalent to:

$$(PNC_{4b}) \quad \neg(\vdash_x \alpha \wedge \vdash_x \neg\alpha)$$

This brief overview on the different versions of the principle of non-contradiction that can be found out in Aristotle's works is useful to better understand what is the attitude of the dialetheist towards it. In the words of Priest:

²⁶ In LP we have that $\neg T(\alpha) \rightarrow T(\neg\alpha)$, but we don't have the other way round, that is, $T(\neg\alpha) \rightarrow \neg T(\alpha)$. Hence, in LP, (PNC_{2b}) is not equivalent to (PNC_{2c}). On the contrary, the equivalence is valid both in Field and Beall's theories. I'll come back to this point in more details in chapter 4.

²⁷ Both Priest and other non-classical logicians – such as Hartry Field, who develops a paracomplete theory of truth – endorse a non-classical account of acceptance and rejection. For the details see section 1.3.2.

Dialetheism

It is sometimes said that dialetheism is a position based on sand. In fact, I think, is quite the opposite: it is the LNC [i.e. Law of Non-Contradiction] that is based on sand. It appears to have no rational basis; and the historical adherence to it simply dogma. Hence it fails. [Priest 1998, p.419]

Hence, on the one hand, it appears very clear from this quote that there is one sense in which Priest disputes the principle of non-contradiction. On the other hand, instead, despite he throws down the gauntlet to the Principle, we will see that Priest doesn't completely reject any of its interpretations.

In fact, counter to expectations, Priest actually accepts all the versions of the principle of non-contradiction pointed out by Łukasiewicz. Indeed, as things are regarding the relation between LP and classical logic, we should have expected this sort of result. It might be useful to remember that classical logic is nothing but a weaker sub-logic of LP, because the latter looks exactly like the former except for few rules of classical logic that are not valid in LP. In particular, classical logic is a special case of LP in which the truth-values are restricted to "true" and "false". Since the semantics of LP subsumes the one of classical logic, it will also include all logical truths of classical logic, one of which is exactly the law of non-contradiction. For this reason, all formulations of the principle of non-contradiction are logically true in LP.

Unfortunately, things are not so easy, because if, on the one hand, it is true that all versions of the Principle are true in LP, on the other hand, it is also true that some counterexamples to them are true as well in the logic of paradox. This is the real sense in which the dialetheist challenges the principle of non-contradiction: according to the spirit of dialetheism, in LP we have that both the law of non-contradiction and certain of its counterexamples are true. For instance, the logico-semantic formulation of the principle (i.e. that a sentence cannot be both true and false) is logically valid in LP, but we also have that the liar sentence (a sentence that says of itself that is false) is both true and false in it. The same goes for the other formulations of (PNC).

Actually, there is an exception to this dialethic attitude. In fact, there is a formulation of the principle that doesn't admit counterexamples even for a dialetheist logician. It's the first psychological-pragmatic version of the principle of non-contradiction, i.e. (PNC_{4a}), which holds that one can't both accept and reject the same sentence at the same time. The reason is due to the non-standard account of acceptance and rejection endorsed by the dialetheist, according to which the two mental states of acceptance and rejection are exclusive and, so, cannot be compatible at all. In this case, we have that (PNC_{4a}) is valid in LP and no counterexample can be provided as well.²⁸

To sum up, all the formulations of the principle of non-contradiction are valid in LP, but this doesn't mean that dialetheists accept the principle in a classical acceptable way. As we have seen, they still throw down the gauntlet to the *principium firmissimum* in a way that is perfectly coherent with the dialethic approach, namely by accepting both the principle and its negation. We can conclude with Priest's words that briefly summarize the meaning of this section:

²⁸ For the details of the dialethic account of acceptance and rejection, see section 1.3.2.

I am attacking the law of non-contradiction, which has been a part of all logical theories. In a sense, this is true. I am attacking the law of non-contradiction (though I am, of course, prepared to assert it too) in a way that is unthinkable for a classical, Aristotelean, or even intuitionist logician. [Priest 1987, p.208]

1.3 Acceptance and Rejection

1.3.1 The Traditional View

Acceptance and rejection belong to the domain of pragmatics and are traditionally and almost universally defined as cognitive states. This means that accepting something is nothing but believing it, and rejecting something seems to be refusing to believe it. The speech acts corresponding to acceptance and rejection are respectively assertion and denial. The latter are closely related to their own cognitive states: in fact, asserting something indicates that one is accepting that thing, and, denying something expresses the fact that one is rejecting it. Hence, the speech acts used by a person reveal her corresponding cognitive states.

According to the classical account, acceptance and rejection are exclusive but non-exhaustive. They are exclusive because we can't accept and reject something at the same time. They are non-exhaustive because agnosticism is a third possible way: for instance, one is allowed not to accept something because she has never taken it into account, so she is refraining from accepting it, but this doesn't mean that she is rejecting it; she can simply remain neutral with respect to the issue, that is, she can neither accept nor reject it.

Furthermore, according to classical logic, acceptance and rejection are two interdefinable notions. In fact, the rejection of a claim is equivalent to the acceptance of the negation of that claim. It's because of this equivalence that we have the convergence between the two versions of the psychological-pragmatics formulation of the principle of non-contradiction, i.e. (PNC_{4a}) and (PNC_{4b}).

1.3.2 The Dialetheic View

As we can imagine, things are quite different for dialetheism. In particular, Priest rejects the interdefinability of acceptance and rejection. The reason is that there are some situations where we can accept both a sentence and its negation. However, this doesn't have the awkward consequence of both accepting and rejecting that sentence. In other words, according to the dialetheist, we can accept $\neg A$ without being forced to reject A .

Therefore, also for the dialetheist, acceptance and rejection are two non-exhaustive but exclusive notions. On the one hand, they are non-exhaustive for a different reason from that of classical logic: in addition to cases of agnosticism, there are other peculiar situations – namely, dialetheias – where we have to accept both a sentence and its negation, such as, for instance. This is the case of the liar sentence. We are allowed to accept both the liar and its negation. On the other hand, acceptance and rejection are exclusive, that is,

they are incompatible notions, because, also according to the dialetheist, we can't in any case accept and reject the same sentence at the same time, and in particular:

If one is presented with a claim that is *prima facie* both rationally acceptable and rejectable, this conflict must be resolved in favour of one or other party, or of agnosticism. [Priest 1987, p.103]

Therefore, rejection has to be kept as a primitive notion, not dependent on the notion of acceptance. Therefore, the two notions are dual to each other. We can better explain this idea using the words of Field:

Rather than explaining rejection in terms of acceptance (as in commitment not to accept), we should regard acceptance and rejection as dual notions. And how exactly one thinks of rejection will depend on how one thinks of the dual notion of acceptance. [Field 2008, p.74]

This quote is very important because allows us to open the issue of rejection and acceptance also to another non-classical theory of truth – i.e. the paracomplete theory developed by Hartry Field – in order to highlights its behaviour towards this matter. In fact, as it turns out from the quote above, even Field has a non-classical account of acceptance and rejection. Also according to him, the two notions are not interdefinable, but in a pretty different way from Priest's one: according to Field, the rejection of a claim is not equivalent to the acceptance of its negation, because we can reject a sentence A without being forced to accept its negation. This means that there are situations where we can reject both a sentence and its negation.²⁹ Field's position on this matter is dual to Priest's one: in fact, while Priest accept only one side of the equivalence between rejecting a sentence and accepting its negation ($\neg_x \alpha \leftrightarrow \vdash_x \neg \alpha$) – i.e. the one from left to right – and reject the other one; Field, on the contrary, does exactly the opposite – i.e. accept the right to left side and reject the other.

To conclude, it is worth mentioning the main advantage of this account for a dialethic logician. As we will see in section 1.5.2, taking rejection as a primitive notion allows the dialetheist to have a notion capable of expressing the disagreement – that is one of the main problems, if not perhaps the only problem, for a dialethic theory of truth –, even if this is not a completely satisfactory solution because rejection is a pragmatic and not a semantic notion.

1.4 The Liar

1.4.1 Logical Paradoxes

Logical paradoxes are paradoxes of self-reference,³⁰ whose definition is provided by Priest:

²⁹ Field[2008], p.74.

³⁰ The so-called paradoxes of self-reference are also known as paradoxes of circularity.

The paradoxes are all arguments starting with apparently analytic principles concerning truth, membership, etc., and proceeding via apparently valid reasoning to a conclusion of the form “ α and not- α ”.
[Priest 1987, p.9]

The logical paradoxes of self-reference are exactly one of the reasons adduced in favour of dialetheism. They can be divided into two different categories: the category of semantic paradoxes, which involve semantic notions, such as “truth”, “predication”, “denotation”, etc.; and the category of set-theoretic paradoxes, which involve set-theoretic notions, such as “cardinality”, “membership”, etc. These two categories have been considered as two sharply distinct phenomena for many years in the past, and, as such, they required distinct solutions. The reason of this distinction is due to the English philosopher and mathematician Frank Ramsey, who claims in 1931 that semantic paradoxes, contrary to set-theoretic ones, give rise to not-purely logical contradictions, because they involve empirical notions.³¹ More recently, we could see through this particular fanciful distinction, thanks especially to the methods of formalization introduced by Gödel and Tarski, who have the merits of having introduced procedures for obtaining non-contextual self-reference.

I’m not going through set-theoretic paradoxes here. I will only deal with semantic paradoxes and, in particular, I will focus on a specific semantic paradox, known as the liar paradox. After all, Priest himself centres his analyses on the liar, which is one of the main motivations for his dialethic theory:

Although cases for the existence of dialetheias can be derived from almost any paradox of self reference, we will focus only on the Liar, given that it is the most easily understandable and its exposition requires no particular technicalities. [Priest and Berto 2013]

1.4.2 The Liar Paradox

The liar paradox is a very old and famous self-reference paradox. Despite it has been studying for more than 2000 years, no satisfying solution has been found yet. The liar is a sentence that says of itself that it is false and it can be formalized in the following way:

$$(L) \quad F(\langle L \rangle)$$

where F is the falsity predicate and $\langle L \rangle$ is a name for (L) . Why is this sentence paradoxical? Let’s reason by cases. First, if we assume that (L) is true, then, what it says is true, but it says that it is false, so, if (L) is true then is false. Secondly, if we assume (L) to be false, then this is exactly what (L) says and hence it will be true. So, if (L) is false then is true. Hence, as a result, we have that (L) is true iff (L) is false. This is equivalent to the conjunction of “ (L) is true” and “ (L) is false”. Hence, (L) is a paradox.

³¹ Ramsey[1931].

Furthermore, the liar is particularly problematic for those theories that aim at keeping the T-schema safe. In order to provide an inductive definition of truth, the Polish mathematician and logician Alfred Tarski states the so-called T-schema, which every theory of truth should be subject to.³² Presented in this way, the T-schema is an intuitive desideratum for quite all theories of truth and its formalization is the following:

$$(T) \quad T(\langle a \rangle) \leftrightarrow a$$

where T is the truth predicate and $\langle a \rangle$ is a name for the sentence a. From the liar, the T-schema and by Substitution of Equivalents,³³ we obtain the following:

$$T(\langle L \rangle) \leftrightarrow F(\langle L \rangle)$$

That is, an implicit contradiction: the liar is true iff it is false. To make it explicit we have that the liar is true and the liar is false.

According to Priest, no satisfying solution to this problem has been found yet.³⁴ Of course there are some options available – and we will see them in next sections – but none of them provide an acceptable result. The reason is well explained by Priest himself:

A solution [to paradoxes] would tell us which premise is false or which step invalid; but, moreover it would give us an independent reason for believing the premise or the step to be wrong. [Priest 1979, p.220]

If it is not possible to find out this reason, then the solution is *ad hoc*, and so unsatisfying. According to the English philosopher, all the existent attempts to find a suitable solution to the liar paradox fail this test and, hence, the problem still remains open.

To conclude, in this section I've shown that the liar, along with the T-schema, leads to inconsistency. Which possibilities do we have to solve this problem?

1.4.3 Classical Theories vs Non-Classical Theories

There are four possibilities for a theory of truth in order to avoid or solve the problem due to the liar:

- A. To give up the notion of truth that obeys the T-schema and to substitute it with one that doesn't.
- B. To keep a notion of truth that obeys the T-schema, and, on the other hand, prevents the construction of the liar by banning the self-reference.

³² I'll give a more detailed analysis of the T-schema and the importance of Tarski's contribution to formal semantics in section 2.1.2.

³³ L is equivalent to T($\langle L \rangle$) by construction.

³⁴ Priest[1979].

- C. To restrict the T-schema in order to exclude the liar from the acceptable instances.
- D. To restrict classical logic by rejecting one or another of its rules.

Let's now have a look at each of these four possibilities.

With regard to A – that says that the liar asks for a different use of the truth predicate from the one used into the paradox –, it corresponds to a particular kind of theories, known as contextual theories. Hartry Field satisfactorily sets them out:

By a contextual, or indexical, theory of truth I mean one according to which the content of any individual sentence-token involving “true” depends upon the context of utterance, where this includes what others have recently said in the conversation, and many perhaps include also the speaker’s intentions. [Field 2008, p.211]³⁵

This means that, according to contextual theories, the liar is context-dependent. The problem with this account is that it is subject to a strengthened version of the paradox.

The second possibility, i.e. B, proposes a way to escape the construction of the liar by forbidding the self-reference. The problem with this purpose is twofold: on the one hand, there are some forms of self-reference – for instance, the Gödel-Tarski diagonalization –³⁶ that we cannot remove without severely limit the expressive power of the language; on the other hand, we can always create new paradoxes, very similar to the liar one, that don't require any form of explicit or direct self-reference.³⁷ Therefore, also B is not a satisfactory solution to the problem raised by the liar.

Let's consider now the two remaining possibilities, i.e. C and D, that are the most interesting for the present discussion. They correspond to two diametrically opposed moves we can choose from: on the one hand, if we choose C, we decide to keep unchanged classical logic by restricting the T-schema; on the other hand, with D, we decide to keep the full T-schema safe and, by contrast, to restrict one or another rule of classical logic. There are advantages in both choices. Validating classical logic has the advantage of simplicity and familiarity, but, as we will see below, it has some important and awkward disadvantages. On the other hand, by weakening classical logic we clearly cannot enjoy the comfort of familiar reasoning, but, at the same time,

³⁵ Examples of contextual theories of truth are Parsons[1974], Burge[1979], Simmons[1993] and Glanzberg[2004].

³⁶ The *Diagonalization Lemma* states the existence of self-referential sentences in certain formal theories of natural numbers.

³⁷ For instance, this is the case of paradoxes involving two sentences, each referring to the other:

- (1) (2) is false.
- (2) (1) is true.

If we assume the (1) is true, then (2) is false for what (1) says. But if (2) is false, then (1) turns out to be false, against our assumption. Hence, contradiction. Otherwise, if we assume that (1) is false, then (2) is true. But, if (2) is true, then (1) turns out to be true for what (2) says, against our second assumption. Contradiction again. Both alternatives lead to contradiction, so we are dealing with a paradox. Therefore, this is an example of paradox without any direct form of self-reference.

we can keep the full T-schema safe without any restriction, which is a very important desideratum for a theory of truth. Whatever we decide, if we want to solve the liar paradox, and the semantic paradoxes in general, we have to give up something in favour of something else: with classical logic we have to “change”, in some way, the meaning of the truth predicate because we can’t have all the Tarskian biconditionals, that seem to be a *sine qua non* of any theory of truth; on the other hand, with non-classical logics, one can say, we have to “change” the meaning of negation, because we are forced to reject some rules governing the use of this connective.

Let’s now individually analyse each of these possibilities. Let’s begin with C, according to which, in order to preserve classical logic we have to restrict the T-schema, that is, we have to reject certain of its instances. Let’s follow the analysis provided by Field and consider three universally accepted principles about incoherence:

- *First Incoherence Principle*: it is incoherent to accept both A and $\neg T(\langle A \rangle)$.
- *Second Incoherence Principle*: it is incoherent to accept both $\neg A$ and $T(\langle A \rangle)$.
- *Third Incoherence Principle*: if accepting either of two claims is incoherent, then it is incoherent also accepting their disjunction. This means that if it is incoherent to accept each of the following conjunctions, $A \wedge \neg T(\langle A \rangle)$ and $\neg A \wedge T(\langle A \rangle)$, then it is incoherent to accept also their disjunction: $(A \wedge \neg T(\langle A \rangle)) \vee (\neg A \wedge T(\langle A \rangle))$.

In order to solve the liar, this strategy forces us to reject its correspondent instance of the T-schema. In particular, we have to reject the instance of the T-schema for the liar. Since, in classical logic rejection is equivalent to acceptance of negation, to reject the instance of the T-schema for the liar means that we accept what follows:

$$(TL) \quad \neg(T(\langle L \rangle) \leftrightarrow L)$$

Since, the biconditional in this sentence is the material one, then (TL) is equivalent to the following disjunction:

$$(TL^*) \quad \text{Either } (T(\langle L \rangle) \wedge \neg L) \text{ or } (L \wedge \neg T(\langle L \rangle))$$

This means that we have three possibilities to deal with the liar if we assume this classical solution: we can accept the first disjunct, we can accept the second disjunct or we can assume an agnostic position between them. The situation, hence, is that we have to accept one of the following three options:

1. $T(\langle L \rangle) \wedge \neg L$;
2. $L \wedge \neg T(\langle L \rangle)$;
3. $(T(\langle L \rangle) \wedge \neg L) \vee (L \wedge \neg T(\langle L \rangle))$.

As one would expect, this is a problem for classical theories. As a matter of fact, the first conjunction is a clear violation of the second incoherence principle, and the second conjunction violates the first incoherence principle. At this point one could think that a third way is possible because it seems that nothing violates the third principle of incoherence. Unfortunately this is not the case. The third possibility is the disjunction of the two statements and corresponds to an agnostic position between them. For this reason, it violates precisely the third incoherence principle, because it corresponds exactly to the situation of a disjunction of two incoherent claims. Therefore, to conclude, we have seen that, in any case, also by restricting the T-schema what we obtain is incoherence. The only way you have to preserve classical logic is by violating some of these three principles about incoherence and this is a very awkward result for a theory of truth.³⁸

At this point, the only option left is D, namely, restricting some rules of classical logic in order to keep the full T-schema. There are two ways in which we can achieve it: by rejecting the excluded middle or by rejecting the explosion rule. The former is the case of paracomplete theories, and the latter is the case of paraconsistent ones. In the present chapter we are interested in the last one, in particular in the solution to the liar provided by a specific paraconsistent theory, i.e. dialetheism. However, in chapter 4, I'll consider one kind of paracomplete theory, the one developed by Hartry Field, and the solution he gives to this paradox.

1.4.4 The Dialethic Solution

After having shown that the first three possibilities fail in their attempts to solve the problem of semantic paradoxes, what is left is the analysis of the last alternative, i.e. D, and, in particular, of the solution provided by dialetheists. According to Priest,³⁹ the only satisfying solution is to admit paradoxes as *brute* facts, or, in other words, to admit inconsistency as a *brute* fact and semantic paradoxes as a proof of this claim. So, let's see in more detailed how the dialetheist tries to solve the liar paradox.

The liar, remember, is a sentence that says of itself that it is false:

$$(L) \quad F(L)^{40}$$

where F is the falsity predicate. If we assume (L) to be true, then it is false for what it says. Otherwise, if we assume (L) to be false, then this is exactly what (L) says, so it turns out to be true. Hence, both alternatives lead to contradiction.

Let's see now which is the analysis of the liar provided by the dialetheist. As we can imagine, according to dialetheists, the liar is a paradox leading to a contradiction that is not only apparent, but real in all respect. This means that

³⁸ Field[2008], pp.119-120.

³⁹ Priest[1979], p.220.

⁴⁰ As we saw in the previous sections, the right formulation would be $F(\langle L \rangle)$ instead of $F(L)$, because we are naming the sentence and not using it. I will often use the simplified version without angle brackets only for convenience, but it's important to keep in mind that we are always talking of the name of a sentence and not of the sentence itself.

the answer given by a dialetheist, such as Priest, to the issue raised by (L) is that it must be evaluated with a third different truth-value, i.e. different from “true” and “false”, and corresponding to “both true and false”. Furthermore, the liar demonstrates that there are proper dialetheias, and that our attitude towards them must be of fully acceptance. Indeed, we should accept these true contradictions and, in order to do so, we should modify our logic so that trivialism doesn’t follow. The only way to achieve this goal for a dialetheist is by leaning to a paraconsistent logic.

A first objection to this approach immediately arises: the theory turns out to be self-defeating. We can formalize the solution to the liar provided by the dialetheist in this way:

(S) (L) is true and (L) is false

Which is the truth-value of (S)? If (L) is both true and false, also “(L) is true” and “(L) is false” have the same truth-value. But, by the dialethic truth table for conjunction, if two conjuncts are both true and false, so is their conjunction. Hence, (S) is true and false at the same time, and, in particular, (S) is false. Let’s take the argument a step forward. According to the dialethic truth table for negation, if (S) is both true and false, then the negation of (S) has the same truth-value. In particular, this means that the negation of (S) is true. From this result, along with the T-schema, it follows that it is not the case that (L) is true and (L) is false. But this is exactly the negation of (S), which, remember, is the solution to the liar given by dialetheists. In this way, it seems that dialetheists are negating what they are affirming. The theory seems to be contradictory itself, because from the theory it follows that the very core of the theory – namely, the dialethic claim about the liar – is false.⁴¹

Obviously, Priest has taken into account this objection and answered to it in a very intuitive way: on the one hand, it is correct that (S) is clearly false; but, on the other hand, (S) is also true, and that’s enough for him. As a matter of fact, according to Priest, if a sentence is true, it must be accepted because true is the *telos* of assertion; on the contrary, if a sentence is false, then this is not a sufficient reason to reject it, because we could have an agnostic attitude towards it or, alternatively, the sentence could be true as well and hence we should accept it, as we have already seen in section 1.3.

To conclude this section, it seems that Priest has found a satisfactory way to solve the problem of semantic paradoxes, i.e. they are both true and false. In other words, according to the dialetheist, we should take inconsistency as nothing but a hard fact of life, and consider the liar as an exemplification of this. In fact, according to Priest, inconsistency “is just an unfortunate fact of life”⁴² and “there is nothing a classical logician can say to force consistency.”⁴³ In next section, we will see that it’s not all peace and light for the dialetheist.

⁴¹ For the details of the objection see Littmann and Simmon[2004].

⁴² Priest[1987], p.106.

⁴³ Ibid., p.112.

1.5 Problematic Aspects

1.5.1 Revenge

We have just seen which is the dialethic behaviour with respect to the liar paradox. Unfortunately, problems related to the liar are far from being solved in this way. In fact, in this section we have to deal with a new and strengthened version of this paradox, commonly known as Revenge Liar.

The revenge phenomenon is a particular reformulation of the semantic paradoxes, which involves the key notions of the solutions to those paradoxes provided by different theories. As we can imagine, this strengthened and more sophisticated version of the liar paradox is called revenge liar because it seems to revenge the standard one by rising from its axes. It is important to highlight that different theories will face different versions of the revenge liar depending on the notions they use to solve the standard paradox.

According to Priest, the revenge phenomenon is easy to explain. He proposes to consider the totality of sentences as divided into two subsets: the subset of true sentences, called *Good Guys*, and its complement, called *the Rest*. According to the dialetheist the essence of the liar paradox can be summarized as follows:

A particular twisted construction which forces sentences, if it is in the *bona fide* truths, to be in the Rest (too); conversely, if it is in the Rest, it is in the *bona fide* truths. The pristine liar ‘This sentence is false’ is only a manifestation of this problem arrived at by taking the Rest to be the false. [Priest 1987, p.23 (emphasis in original)]

In other words, in a nutshell, the standard liar and the revenge liar are nothing but two different manifestations of the same phenomenon. According to this view, we have the standard liar when the theoretical framework is the standard one – i.e. a 2-valued semantics, that is, when we have only two truth-values, corresponding to “true” and “false” – and we have the revenge liar when the theoretical framework is extended with the new value used to provide a solution to the standard liar. This shows that the revenge liar is not a new paradox sharply different from the standard one, but it only represents the same problem applied to different theoretical frameworks. Therefore, if we try to work around the problem by reformulating the liar with categories that are not in the Rest, then the original problem simply arises again, because the Rest can be easily rewritten in such a way to include also the new categories. To conclude the point, this means that the revenge liar is only a different formulation of the same problem and nothing else. However, this doesn’t weaken the power of the revenge liar that turns out to be a real problem even for a dialethic theory.

After having explained what the revenge problem is according to Priest, what is left is checking whether the problem can take place even for dialetheism, or not. As we can image, Priest claims his theory to be immune from the revenge phenomenon. According to the dialetheist philosopher the problem affects only those theories – both classical and non-classical –, which

have consistency among their *desiderata*. To be more specific, we can use Priest's words one more time:

A standard objection to proposed consistent solutions to the semantic paradoxes is that they all seem vulnerable to "revenge" paradoxes. There is a certain notion the intelligibility of which the theorist presupposes which, if it is included in the language in question, can be used to refashion the paradox. Hence consistency can be maintained only at the cost of incompleteness – which naturally gives rise to a hierarchy of metalanguages, and so to familiar problems of the same kind. [Priest 2005, p.44]

In other words, the problem is that all the accounts that wish to preserve consistency are inevitably vulnerable to a revenge liar, and the only possible solution available to them is denying some key notions – i.e. exactly the notions used in the formulation of the revenge liar – to be indeed expressible in the language of their theories. This is a very high price to pay for these theories, because asserting that the notions involved in the formulation of the revenge liar are not expressible in the language of the theories themselves means that those paradox-solvers are talking in a language different from the one for which the semantics is offered.

However, Priest asserts that this very awkward consequence doesn't obtain for those theories, such as dialetheism, for which consistency is not a binding requirement. Not everyone agrees with this last claim. More than one who disagrees can be found out in specialist literature. Most of the commentators maintain that also dialetheism suffers the problem of the revenge liar.

Let us take one thing at a time and, to begin, have a look at a first possible candidate to play the role of revenge liar for a dialethic theory of truth:

(R) (R) is false only

Let's now try to evaluate (R). If we assume (R) to be true, then, for what it says, (R) turns out to be false only. But it seems natural to think that (R) cannot be both true and false only, because of the meaning of the adverb "only". Hence contradiction. Suppose now that (R) is false only, but this is exactly what (R) says, hence (R) is true. For the same reason above, (R) cannot be both true and false only at the same time. Contradiction again. Both the alternatives lead to contradiction, so we have a paradox. In both cases we have a contradiction that seems not to be solvable even in a dialethic theory of truth.

According to Priest, things are different. He claims he has a satisfactory and intuitive solution to the revenge liar: a dialetheist can simply say that (R) is a dialetheia. Specifically, the dialetheist's answer to the revenge liar above is that (R) is both true and false, and false only. But when we use the value "false only" in the formulation of (R) we want to convey the fact that a sentence with this value cannot be true as well. This is exactly the meaning we want for "false only" and this is also the meaning the adverb "only" commonly have in our language. So, how is it possible for a sentence to be false only and true as well? Priest says that "false only" is not to be understood in a truth-exclusive sense, contrary to what our intuitions suggest. According to him, we can treat "false

only” in the same way as the value “false”, there is no difference between them. In this way, we can always add a designated value, such as “true” or “both true and false”, to a sentence that is false only. In conclusion, according to dialetheism, nothing precludes a sentence that is false only to be true as well.

A first problem with this solution immediately arises. In fact, it seems that Priest in what he says about the non-exclusiveness of “false only” is violating the function that the adverb “only” has in the language as a universal quantifier and, as a consequence, the real meaning of the truth-value “false only”. As a matter of fact, when we say that a sentence is false only, we want to exclude all others truth-values from the evaluation of that sentence.

This is the first problem the dialethic solution faces. But there is another and more problematic issue the dialetheist has to deal with. By handling “false only” in the way Priest does, he appears not to be able to express effectively that a sentence is false without being in anyway true. Hence, the problem of the revenge phenomenon is strictly linked to another and more pregnant issue, namely, the expressive limitedness of the language of the theory. I’ll come into a sharper focus with this issue in next section.

I’m going to deal here with other types of revenge liars for a dialethic theory of truth. One of the most recognizable revenge liars for dialetheism is the one suggested by the two philosophers Greg Littmann and Keith Simmons. They propose to consider a sentence free of any contextual reference that we know for sure is false. For instance, let’s consider a mathematical sentence, such as the following:

$$(M) \quad 1+1=3.$$

Even the dialetheist will admit that (M) is false and there’s no way it can be also true. The underlying idea is that arithmetic cannot contain dialetheias, because it is entirely consistent. Littmann and Simmons suggest to name the truth-value ascribed to this sentence “v”. According to what we have just said, if a sentence has the value “v”, then it cannot have any other designated value, because is the value we assign to an arithmetical falsity. Take a look now at the following revenge liar for dialetheism, called by the two philosophers Introspective Liar:

$$(Z) \quad (Z) \text{ is } v.$$

It’s easy to see that a contradiction immediately follows from it. Let’s try to evaluate (Z). If we suppose (Z) to be true, than, for what it says, (Z) turns out to be v. But v is the truth-value corresponding to (M), an arithmetic falsity so, it seems that a sentence cannot be both true and v. Hence, contradiction. On the contrary, if we assume (Z) to be v, this is exactly what (Z) says, so it turns out to be true. But, as we have just said above, v is a truth-excluding value. Contradiction again. In both cases we have a genuine contradiction, so (Z) is paradoxical.

According to Littmann and Simmons, the dialetheist cannot solve this paradox in the usual way, i.e. by saying it’s a dialetheia. For this reason, they claim they found a proper revenge liar for dialetheism.

Actually, even in this case, Priest has an answer and it corresponds to the usual answer a dialetheist provides to paradoxical sentences. As a matter of fact,

he claims he can coherently solve the paradox by saying that (Z) has the following semantic status: “both true and v”. But this means that there is a sentence that has the same truth-value of “ $1+1=3$ ” – i.e. an arithmetic falsity – and that is true at the same time, and it’s not clear how this can happen. In particular, according to Littmann and Simmons:

If the dialetheist ascribes truth to a sentence that is v, then it seems to us that she has failed to understand what it is for a sentence to be v. Add ‘true’ to ‘v’, and you no longer have ‘v’. [Littmann and Simmons 2004, p.324]

To conclude, in this section I’ve tried to give an overview over one of the main issues dialetheism has to deal with: the revenge phenomenon of the liar sentence. I’ve introduced only two possible revenge liars for dialetheism and the correspondent solutions proposed by Priest, but there are other attempts to provide a genuine revenge liar for dialetheism, that is, a detonator that would cause dialetheism to explode.⁴⁴

In any case, even if dialetheism could properly face the problem of the revenge phenomenon and could correctly solve each of the revenge liars that would arise, another issue would still remain open, i.e. the expressive limitedness of the language, which is strictly connected to the revenge phenomenon and, in particular, to the solution to the revenge liar provided by the dialetheist.

1.5.2 The Expressibility Problem

The problem of the expressibility of the language of a theory is a very important problem affecting most of the theories of truth that aim at solving the liar – and, perhaps, it affects all of them, as we will see. In fact, the problem of the expressibility arises directly from the solution a theory provides to the semantic paradoxes, in particular to the liar. As Priest pointed out, this is commonly a problem for those theories of truth that wish to preserve consistency:

The only way that various consistent accounts of the paradoxes can be maintained is by taking some notion, to the legitimacy of which the proponent of the solution is committed, not to be expressible in the language of the paradox. [Priest 1987, p.291]

By so saying, Priest implicitly claims that only consistent theories face this issue, and not the inconsistent ones, such as dialetheism. By contrast, in literature can be found out several criticisms towards dialetheism that move in this direction.

The problem of the expressibility of the language for a dialetheist can be given in many different ways. Some have individuated the problem in the impossibility for a dialetheist to express the disagreement;⁴⁵ according to some others, instead, the problem lies in the inability for a dialethic theory of truth to keep using the instrument of confutation, which has to be understood in an

⁴⁴ See Littmann and Simmons[2004] for a graphic liar and Shapiro[2004] for the irrationality liar.

⁴⁵ See Priest[1987], Murzi and Carrara[2014] and Ripley[2014].

Aristotelian sense;⁴⁶ according to still others, the core of the problem is that the dialetheist is not able to express within the language of its theory the exclusion.⁴⁷

Whatever the matter is meant, the problem for a dialetheist is that of an expressive limitedness of its language. More specifically, the dialetheist seems not to be able to rule out a sentence A by simply saying “ $\neg A$ ” – as we usually do –, since A could be a dialetheia. In fact, if A is a dialetheia, we will have both A and $\neg A$. This means that the use of negation “ \neg ” in a dialethic logic is no more guarantee of exclusion of what is negated. The same goes for “ A is false”. Even in this case, when a dialetheist says that A is false, in no way this prevents A from being true as well, because in a dialethic logic “ A is false” does not automatically rule out A ’s being true as well – contrary to what happens in classical logic. In other words, it seems impossible to express the idea that a sentence is true only or false only – that is, the idea that a sentence is not a dialetheia – within the dialethic language itself. The reason is that the non-designated truth-values “false” and “false only” are logically compatible with the designated ones i.e. “true” and paradoxical. Hence, as a consequence we have that if one asserts a sentence, we are not able to express disagreement with that sentence using the dialethic language.

Priest claims he actually has a way to express that a sentence is not a dialetheia. In fact, he can express that a sentence is not a dialetheia precisely by using those very words, i.e. by saying that it is “false and not true”, or “false and not a dialetheia” (the same also goes for “true and not false” and “true and not a dialetheia”). The point stressed by Priest is that he can express that a sentence is not a dialetheia in this way but he can’t do it in a consistent way. However, this is not a problem for a dialetheist and, in fact, Priest remember us that:

There is no statement that can be made which forces α to behave consistently. This is one of the hard facts of dialethic life. Actually, it is one of the hard facts of life, *period*. There is nothing a classical logician can say to force consistency either, and any attempt that fails will occasion an immediate collapse into triviality, the highest degree of inconsistency, rather than merely a higher degree of inconsistency, as in dialetheism. [Priest 1987, p.112 (emphasis in original)]

This does not seem a satisfactory answer to the problem of the expressibility of the language, but rather it seems to evade it and simply to reformulate it in other words. Field explains very clearly which is the real problem in the answer provided by Priest:

The problem that these authors are raising is that on these definitions, the notions don’t behave in accordance with how they seemed intended to behave when the theory was being explained. [Field 2008, p.386]

⁴⁶ See Berti[2014].

⁴⁷ See Berto[2006, 2007 and 2014].

Therefore, “false only” and “true only” can be defined in the way Priest does, but this does not solve in any way the problem of the expressive limitedness of the language.

One move that is still available to the dialetheist is to express the disagreement not in a semantic way – because we have just seen that is no more possible in a dialethic theory –, but in a pragmatic way, by using the notion of “rejection” and its corresponding speech act, i.e. “denial”, that we have already met and analysed in section 1.3. In this way, when the dialetheist wishes to express her own disagreement with a sentence β , she won’t say anymore that β is false only – or, alternatively, that is not the case of β –, but she will simply deny β . Since, according to the dialetheist as well as the classical logician, rejection and acceptance are mutually exclusive, we will be sure that when a dialetheist rejects a sentence, he is not accepting it at the same time, and so we will be sure that that sentence is not a dialetheia. On the other hand, when a dialetheist is dealing with a dialetheia, she will accept both it and its negation, but she won’t reject it in any case.

This approach has a great advantage that “false only” doesn’t have, namely, it is not subject to the revenge phenomenon of the liar. The reason is that the pragmatic operator (that is not a logical one) corresponding to denial is a force-operator who attaches to the whole sentence and “has no interaction with the content of what is uttered.”⁴⁸

Despite the advantages, this proposal faces a big deal that disqualifies denial for the role of expressive device for disagreement. This awkward disadvantage is highlighted both by Hartry Field⁴⁹ and, even before, by Stewart Shapiro.⁵⁰ Using “denial” to rule something out turns out to be a very limited solution, because it can’t be used in richer embedded sentences, since it is not able to capture their real meaning. For instance, using an example from Field, it is not clear how to capture in terms of rejection the meaning of a sentence such as “if the premise of a conditional is solely true and the consequent is solely false, then the whole conditional is solely false”. More specifically:

Suppose that Karl says “ β ”, and his dialetheist friend Seymore does not wish to disagree (yet), but he wonders if Karl is mistaken. Seymore might want to assert a conditional in the form: “if Karl is mistaken, then φ ”. How can Seymore express this? Again, “if $\neg\beta$ then φ ” won’t work. Since, for Seymore, $\neg\beta$ is compatible with β , it is not the way for him to say that Karl is mistaken in asserting β . I do not see how implicature helps here. What are the conversational rules for formulating hypotheses, or for the antecedents of conditionals? Even if there are coherent and useful implicatures concerning hypotheses, they cannot be used to determine the consequences of these hypotheses. So far, we just do not have a statement equivalent to “Karl is mistaken in asserting β ”.
[Shapiro 2004, pp. 339-340]

Hence, this approach seems not to be totally satisfying.

⁴⁸ Priest[2006], p.108

⁴⁹ Field[2008]

⁵⁰ Shapiro[2004].

Thankfully, there is another move available to the dialetheist, i.e. the so-called “arrow-*falsum*” strategy.⁵¹ For instance, we can disagree with a sentence β by saying:

$$\beta \rightarrow \perp$$

where \perp is something that a dialetheist rejects (for instance, that everything is true, or that I’m a fried egg, to use an example by Priest). This will allow us to rule β out because since \perp is something we reject, we will reject β as well.

This approach seems to face two types of problems. First, arrow-*falsum* is not a very intuitive way to express disagreement. In fact, as Berto shows:

A dialetheist from Hartford may want to disagree on “Hartford is in Rhode Island” on the basis of plain empirical evidence. It seems strange that he can only express this by claiming, “If Hartford is in Rhode Island, then the absurd *falsum* obtains”. [Berto 2014, p. 196 (emphasis in original)]

Secondly, another and more important problem with this approach is that it leads to triviality because of the Curry sentence, which says that from its truth the absurd *falsum* follows:

$$(K) \quad T(K) \rightarrow \perp$$

(K) can be neither true nor both true and false, because if it had a designated value, it would explode into trivialism. Hence, it must be ruled out. However, we can’t do this by saying “ $(K) \rightarrow \perp$ ”, because by the T-schema and Contraction Rule ($A \rightarrow (A \rightarrow B) \vDash A \rightarrow B$) this sentence is equivalent to (K) itself, and we would have absurd *falsum* by *modus ponens*, and, as consequence, trivialism would follow. So, also arrow-*falsum* is not a satisfying approach.

Actually, there are other attempts to find an appropriate way to solve the problem of the expressive limitedness of the language of a dialethic theory,⁵² but I’ve decided to consider only the ones taken into account by Priest himself, that are also the most known and debated into literature.

To conclude this section we can say that at this point we haven’t found yet a suitable solution to the problem, which, at this level, still remains open. So, from the analyses of the issues faced by a dialethic theory, we can conclude that, for now, only one of the following alternatives is possible: either we solve the revenge liar and, as a result, we decide to severely limit the expressive power of our language; or we preserve the full expressibility of the language, but, by so doing, we must give up the idea of providing a solution to the problem of the semantic paradoxes in their revenge formulation.

⁵¹ Berto[2014], p.196

⁵² For instance, Berto[2014] offers a possible solution to the problem of exclusion for a dialetheist. Berto tries to state a proposal in accordance with the Priestian program, by keeping together the existence of dialetheias with the possibility to express the exclusion of two incompatible properties. Berto’s account is built upon a treatment of negation through the notion of material incompatibility, which should allow explaining why not all contradictions are true, according to dialetheism.

Chapter 2

Deflationism about Truth

2.1 Introduction to Deflationism

There are two possible ways to interpret the following famous Aristotelian passage:

To say of what is that it is not, or what is not that it is, is false, while to say of what is that it is, or of what is not that it is not, is true. [Aristotle 1928, 1011b25]

The first way is to maintain that a sentence is true if it corresponds to something real, i.e. if there is something in the world that makes it true. Adopting this approach means adopting a *Correspondence view* of the nature of truth. The second way to interpret the Aristotelian quote requires a less demanding metaphysical understanding of it, and it corresponds to the philosophical position commonly called *Deflationary view* of the nature of truth.

Despite the former interpretation is the one that is traditionally adopted from the majority of philosophers, in the last decades a great interest and enthusiasm for the latter type of perspective has arisen. The core of this chapter is this second way of understanding truth, that is, a less metaphysically committing way.

2.1.1 Deflationary Theory of Truth in a Nutshell

Deflationary theories of truth follow from the attempt to provide an answer to questions like “What is truth?”, “What is the nature of truth?”, and so on. Therefore, deflationism is a theory about the nature of truth, or rather a theory about its non-nature. As a matter of fact, according to deflationary theories of truth, the truth predicate expresses no genuinely substantive property, if there is any property of truth at all. A property is substantive if it owns an underlying nature.¹ There’s nothing about the nature of truth that goes beyond what is caught when we say, for instance, that “snow is white” if and only if snow is white. Claiming that a sentence is true means nothing more than claiming the sentence itself, because the truth predicate doesn’t add anything to the meaning of the sentence. As Bradley Armour-Garb puts it:

¹ Naturally, if the predicate does not express any property, then characterizing the property as either “substantive” or not is superfluous.

Chapter 2

The theory is deflated because it effectively denies that the truth predicate admits of a standard, or substantive, definition. [Armour-Garb(2013), p.54]

However, when we say that deflationism is a theory about the nature of truth, we are not saying a totally correct claim, because, as we have already seen, according to deflationism, truth has no underlying nature to take into account.² To better understand deflationism, let's consider the following questions: "What is the nature of truth?" and "What is the conceptual role of truth?" The main difference between deflationary theories and their rivals is that while the latter try to provide an answer to both questions, deflationism, by contrast, takes only the second question seriously. Accordingly, we can say that deflationary theories keep in consideration only the conceptual role of truth and how this concept is used in our talk and thought.³ It's exactly for this reason that I said that deflationism is not properly a theory about the nature of truth, that is, as Armour-Garb well explains:

One who adopts T-deflationism⁴ makes a transition from attempting to theorize about what true is to theorizing about what the expression "true" does. [Armour-Garb and Woodbridge(2015), p.116]

This means that what a deflationist cares is to account for the expressive role performed by the truth predicate, instead of theorizing about the nature of truth, because, actually, there is nothing to theorize about. In other words, according to deflationism, in order to better understand a concept we must, first, account for its function. For this reason, in order to understand what truth is, we must, first, understand, what true does, i.e. what its function within the language is. In a deflationary view, after having analysed what the function of truth is, we become aware that there is nothing more we can say about such concept, because it doesn't have any underlying nature to examine and to take into account.

However, the central features of a deflationary theory of truth are not limited to what I said so far. In particular, what Armour-Garb and Woodbridge call the "T-deflationism's core commitment" consists in the conjunction of four fundamental theses: the rule thesis, the property thesis, the term thesis and the concept thesis. I'll consider each of them in details in next sections, but, at the moment, I'm interested only in briefly clarifying what they say. I have already mentioned the rule thesis, which characterizes truth at *semantic* level; more specifically, it holds that the use of a truth-bearer and the use of an ascription of truth to it are intersubstitutable in all transparent contexts.⁵ The

² Cfr. Introduction, footnote 2.

³ Armour-Garb and Beall[2005].

⁴ "T-deflationism" is the abbreviated name Armour-Garb and Woodbridge give to deflationism about truth, in order to make a distinction between it and deflationism about other concepts, such as *reference*, *meaning* and *existence*, and many others. As a matter of fact, we can adopt a deflationary theory of several concepts, not only truth. While deflationists about truth tend to be also deflationists about reference, it is not automatically assumed that that is the rule, namely, that deflationists about truth will be committed to deflationism about all other concepts.

⁵ A context is *transparent* if any two expressions referring to the same thing can be substituted in it *salva veritate*, i.e. without altering the truth-value of what is said. In other words, a context is

property thesis provides a *metaphysical* analysis of the concept of truth, and, as in the case of property thesis, I have already mentioned it above; the property thesis holds that there is no genuine substantive property of truth. The term thesis is set at *pragmatic* level, because it aims at highlighting the function of the truth predicate, namely, the merely expressive role of expressions like “is true” and “is false”. Finally, the concept thesis characterizes the truth predicate at level of *philosophical import*; in fact, it holds that the concept of truth cannot be clarified at a more fundamental and deeper level via the use of other concepts which the truth predicate is connected to.

Before we consider and analyse in details each theses constituting the deflationism’s core commitment, let’s try to acquire some more general elements about the deflationary theories of truth. As we have already seen, deflationism is usually characterized negatively, by specifying what the account rejects – i.e. that truth has an underlying nature –, rather than what it accepts and takes into account. More specifically, the deflationary view of truth is characterized by the acceptance of the role performed by the truth predicate, without being forced to include some metaphysical presuppositions that are commonly associated to such acceptance. In fact, deflationism assures us that the truth predicate has a role – a fundamental expressive role, namely, that of semantic ascent that we will see in section 2.1.3 – and, at the same time, it is not ontologically committed to the property of truth.

In the last one hundred and a half years, deflationism – or, better, some of the main ideas underlying a deflationary theory of truth as it is understood today, namely, a theory endorsing the four fundamental theses that constitute the deflationism’s core commitment – has become a very widespread view in the philosophical panorama and has received a genuine defence by some major philosophers, which include Gottlob Frege, Frank P. Ramsey and Willard Van Orman Quine. As long as Frege is concerned, the little he said about truth suggests that he would have supported a deflationary view. In fact, in the following passages he seems to point to a certain kind of redundancy of the truth predicate:

One can, indeed, say: “The thought that 5 is a prime number is true”.
But closer examination shows that nothing more has been said than in
the simple sentence “5 is a prime number”. [Frege, 1948, p.216]

and, in another paper:

It is also worthy of notice that the sentence “I smell the scent of violets”
has just the same content as the sentence “it is true that I smell the scent
of violets”. So it seems, then, that nothing is added to the thought by my
ascribing to it the property of truth. [Frege, 1956, p.293]

If there are few references leading to deflationism in Frege’s works, things are definitely different according to Ramsey, who is considered the developer of the so-called *Redundancy theory of truth*. The name of the theory is frequently reason of misunderstandings. In fact, the name “the redundancy theory” is

transparent if an expression can be replaced by another with the same reference without changing its truth-value. Otherwise the context is *opaque*.

Chapter 2

often used to point to deflationism itself, but, in many other cases, it is used to name a variety of deflationary theories of truth, i.e. a sub-theory of deflationism. The following quote will very clearly help to state Ramsey's position on this matter:

It is necessary to say something about truth and falsehood, in order to show that there is really no separate problem of truth but merely a linguistic muddle. [...] It is evident that "it is true that Caesar was murdered" means no more than that Caesar was murdered, and "it is false that Caesar was murdered" means that Caesar was not murdered. [Ramsey, 1927, p.157]

Thanks to this passage we can emphasise an important feature of deflationary theories that it has not been mentioned yet. Everything I said about deflationism and what it states with regard to the truth predicate also applies, in the same way, to the falsity predicate. Therefore, in general, all of the four theses constituting the deflationism's core commitment also apply to the falsity predicate: the use of the negation of a truth-bearer – i.e. the use of a falsity-bearer – and the use of an ascription of falsity to that falsity-bearer are intersubstitutable in all transparent contexts; there is no substantive property of falsity, if there is such a property; the falsity predicate, together with expressions like "is false", merely has an expressive role in the language; and, lastly, there are no concepts connected with the falsity predicate that can give us a deeper comprehension and explanation of such predicate.

Quine, similarly to Ramsey, explicitly advocates for deflationism about truth. If Ramsey can be seen as the forerunner, and maybe the founder, of the redundancy theory of truth, similarly, Quine is seen as the developer of what eventually came to be called the *Disquotational theory of truth*. Also disquotationalism is a variety of deflationism and its core ideas can be explained by Quine himself:

The truth predicate is a reminder that, despite a technical ascent to talk of sentences, our eye is on the world. This cancellatory force of the truth predicate is explicit in Tarski's paradigm:

"Snow is white" is true if and only if snow is white.

Quotation marks make all the difference between talking about words and talking about snow. The quotation is a name of a sentence that contains a name, namely "snow", of snow. By calling the sentence true, we call snow white. The truth predicate is a device for disquotation. [Quine, 1970, p.12]

In the above quote, via the introduction of the notion of "disquotation", Quine allows me to briefly mention one of the main features that the truth predicate has according to a deflationary theory of truth, namely, to enable the cancellation of the semantic ascent. This aspect, which will be fully analysed in next sections, is very important because the central thought is that the function of the truth predicate is to serve a specific expressive purpose, namely, the expression of certain infinite conjunctions and disjunctions, and it serves this role in virtue of the fact that it undoes the effect of quotation marks.

In more recent times, the interest towards deflationary theories of truth has grown up and it has intensified. Among deflationism's contemporary advocates I must mention Paul Horwich, Hartry Field and J.C. Beall. In next sections I'm going to state and explain the positions of all these three exponents of deflationism, and, in chapter 4, I'm going to carefully investigate, in particular, the accounts provided by the latter two.

Before I examine the different versions of deflationism, let's now consider each of the four theses that together form the deflationism's core commitment and analyse them in details.

2.1.2 The Rule Thesis: The Equivalence Principle

The rule thesis holds that the use of a truth-bearer and the use of an ascription of truth to it are intersubstitutable in all transparent contexts. This equivalence principle corresponds, at semantic level, to the property called *Transparency of the truth predicate*, that can be achieved by means of the validity of two desiderata: the T-schema and the Intersubstitutivity Principle.⁶ Therefore, to begin with, let's set up and clarify what these desiderata say and why they are so important for a deflationary theory of truth.

In view of that, a brief overview of the work of the Polish Alfred Tarski, one of the most influent logicians and mathematicians – “and perhaps a philosopher of a sort” as he himself pointed out –⁷ of the twentieth century, is needed. The role of the work Tarski did at the beginning of the 1930s about the notion of truth is of great importance to us for a twofold reason: in the first place, the introduction of the criterion of material adequacy, that every theory of truth must comply in order to be a satisfactory theory of truth, is fundamental for the deflationary theories of truth, as we have already said and as we will see in more details; in the second place, it's with the rigour of the Tarskian theory of truth that the problem of semantic paradoxes – and the deflationary theories are far from being safe from it – started to take on a particular importance for all those theories that want to deal with truth in a rigorous and formal way.

The great deal provided by Tarski is to rehabilitate the notion of truth to a more rigorous and mathematical use. With this idea in mind, the aim of the Polish logician was to define the concept of truth in a logically acceptable way. This attempt results in the birth of a new branch of mathematical logic nowadays known with the name of *model theory*, which main aim is the study of formalized languages and their semantic interpretations. As I previously mentioned, Tarski maintains that it is possible to provide an inductive definition of truth for formalized languages, such that it must comply the criterion of material adequacy, imposed by Tarski himself, also known as

⁶ The validity of the T-schema – or, at least, the formulation of the T-schema I gave in section 1.4.2 – is one of the different ways in which it is possible to state the rule thesis, but it is not the only one. Different kinds of deflationary theories give different accounts of what the equivalence principle amounts to, and of the better way we should interpret it. For example, whether the T-schema has to be taken about sentences or, rather, propositions, is part of the discussion. It is also at issue how to interpret the equivalence principle, because different interpretations yield different versions of deflationism. For an overview of the different species of deflationism see section 2.2.

⁷ Tarski[1944], p.369.

Convention T. A definition of truth for a formal language is materially adequate if and only if every sentence of the language which the truth predicate is applied to complies the T-schema – also known with the names of *Tarskian biconditionals* or *equivalence principle* –,⁸ that is, the definition must have, as a consequence, all instances of the following schema:

$$(EP) \quad \langle p \rangle \text{ is true iff } p$$

or, it can be more appropriately formalized in the way we saw in section 1.4.2, that, to recall, is the following:

$$(T) \quad T(\langle p \rangle) \leftrightarrow p$$

where p is a sentence, T is the truth predicate, and the angle brackets $\langle \dots \rangle$ are an appropriate name-forming device,⁹ such that $\langle p \rangle$ is a name for the sentence p . The left-side and right-side of the T-schema are often divided for the sake of convenience:

$$(T\text{-In}) \quad p \rightarrow T(\langle p \rangle)$$

$$(T\text{-Out}) \quad T(\langle p \rangle) \rightarrow p$$

(T-In) and (T-Out) correspond to two inference rules that hold because p is consequence of itself, for all p in the language. The two rules are *T-Introduction* (i.e. introduction of the truth predicate) and *T-Elimination* (i.e. elimination of the truth predicate) and are formalized in the following way:

$$(IT) \quad p \vDash T(\langle p \rangle)$$

$$(ET) \quad T(\langle p \rangle) \vDash p$$

(T-In) and (T-Out) are also known respectively with the names of *Capture* and *Release*. The two names mean to highlight the role of the device of truth, namely, to “capture” and “release” information. For that reason the names tell us what is the behaviour of the truth predicate: in the first case, you can “capture” A with the truth predicate; in the second case, you can “release” A from the truth predicate. Pointing out the two sides of the T-schema is important not only for the reason I just mentioned, but also because classical theorists of truth typically challenge one side or the other of the T-schema in order to provide an adequate solution to the problem of semantic paradoxes.

As we saw above, the T-schema is fundamental for a full understanding of the deflationary approach for two reasons: in the first place, the T-schema is at the root of a deflationary theory of truth; and this, along with the fact that from the T-schema we can easily obtain a contradiction, causes serious problems to

⁸ *The semantic conception of truth* is the name Tarski gives to his own account because it is based on the criterion of material adequacy.

⁹ Other appropriate name-forming devices are, for instances, quotation marks or “the proposition that...”, etc.

deflationism. Let's put the problem of semantic paradoxes off until section 2.3.2, and let's try to explain first why I say that the T-schema is at the root of the deflationary theories of truth.

Deflationism can be defined as the theory according to which the T-schema – or rather its instances – tells us everything we can say about truth, and, in particular, about the nature of truth. In other words, one can fully capture the meaning of the concept of truth only if she is prepared to accept all the instances of the T-schema,¹⁰ and this requirement is both necessary and sufficient for the purpose. In fact, the deflationist, who claims that there is no substantive property of truth, cannot maintain that the instances of the T-schema hold in virtue of the nature of the property of truth. In contrast, they must claim that the T-schema is logically valid.

It is worth highlighting that, despite deflationism is strictly connected with the T-schema and the latter is associated with Tarski, it is far from obvious that Tarski was a deflationist of a sort.¹¹ In fact, the use of the T-schema in the definition of deflationism is an operation due to contemporary deflationary theorists and it is in no way ascribable to the Polish logician.¹²

If what we said so far about the T-schema is enough for a classical theory of truth, things are slightly different for non-classical theories. Let's, now, consider the version of the Intersubstitutivity Principle provided by Field:

Intersubstitutivity Principle: If C and D are alike except that (in some transparent context) one has the formula $A(x)$ where the other has the formula $\langle A \rangle$ is true of x , then C implies D¹³ and vice versa.¹⁴

In other words, for a theory to validate the intersubstitutivity principle, $\text{True}(\langle A \rangle)$ must be everywhere¹⁵ intersubstitutable with A. This means that we can always substitute an instance of “A” with an instance of “A’ is true” in every transparent context.

Contrary to non-classical theories of truth, classical logicians maintain that the T-schema and the principle of intersubstitutivity are equivalent. In fact,

¹⁰ Someone would say that one has to be prepared to accept all the instances of the T-schema that are in no way *defective*.

¹¹ See Stoljar and Damnjanovic[2014].

¹² As a matter of fact, nowadays Tarski's view about the nature of truth is still part of the discussion. The most common idea is that Tarski was inclined to a correspondence theory of truth. The reason is primarily that Tarski considered himself a follower of Aristotle, who, as we know, was the first one in postulating a correspondence between truth and reality, and, for this reason, he is considered the correspondentist philosopher par excellence.

¹³ Of course, the implication is not classical. In fact, in a classical logic we cannot validate the intersubstitutivity principle because of the semantic paradoxes.

¹⁴ Field[2008], p.12.

Another possible way to state the Intersubstitutivity Principle is the following:

Intersubstitutivity Principle:* If D is a sentence resulting from C by substituting any number of occurrences of sub-sentences A for “A’ is true” (in non-opaque contexts), then D and C are either both acceptable, rejectable, or ill-formed.

The two formulations of the intersubstitutivity principle are equivalent, because they lead to the same result, namely, allowing the intersubstitutivity of “A” with “A’ is true” in every transparent context.

¹⁵ In every non-opaque context.

according to non-classical theories we can have the T-schema without the intersubstitutivity principle, and vice versa.¹⁶ Therefore, while a classical theory to have the right features in order to be deflationary is sufficient the full validity of the T-schema, a non-classical theory to be deflationary must validate both the T-schema and the intersubstitutivity principle. Some examples of non-classical deflationary theories of truth are the dialethic theory developed by Beall¹⁷ and the paracomplete one by Field.¹⁸

To summarize and try to clear up the issue, we can say that, with regard to the T-schema and the intersubstitutivity principle, deflationary theories of truth are committed to three fundamental theses. The first thesis claims that applying the truth predicate to something is equivalent to simply saying that thing, and this is equivalent to the acceptance of the principle of intersubstitutivity and of all the instances of the T-schema. The second thesis states that there is nothing else in the meaning of the truth predicate but the T-schema; in other words, everything we can say about the meaning of the notion of “truth” is nothing but the instances of the equivalence principle. Lastly, the third thesis holds that an analysis of the nature of truth is merely an analysis of the meaning of the truth predicate; but, if to give the meaning of the truth predicate is enough to look at the instances of the T-schema, then also the nature of truth can and must do the same.

2.1.3 The Property Thesis: Truth Is Not a Substantive Property

We defined deflationism as the theory according to which the truth predicate doesn’t express any genuine and substantive property. This is the appropriate formulation of the property thesis.

We made also clear that a property is substantive if it has an underlying nature. However, we didn’t exhaust the topic yet. There are few questions that are still open: what does it mean that a predicate doesn’t express any genuine and substantive property? Does it require a sense according to which the truth predicate expresses a property, even if not a substantive one?

Actually, there’s a sense according to which the answer to this latter question is affirmative. To see why take a look, for instance, at the following sentences:

- (1) Rome is the capital of Italy.
- (2) Snow is white.

Even a deflationist would agree that, to some extent, these two sentences seem to share a property, and that this property is truth. After all, both sentences are true and, as a consequence, it’s natural to think that they share the property of truth. Accordingly, there is a sense according to which even a deflationist must

¹⁶ For instance, in LP the T-schema holds, but, on the other hand, it doesn’t validate the intersubstitutivity principle. In contrast, an example of a non-classical theory in which the intersubstitutivity principle holds but the T-schema doesn’t, is KFS, the internal theory of the fixed point. For the details, see Field[2008].

¹⁷ Beall[2009].

¹⁸ Field[2008].

admit that truth is a property and, more specifically, truth is the property that all true things share.

However, on the other hand, to say that there are things that are P is not equivalent to say that those things own a common property, i.e. P. In other words, in order to say that two objects share a specific property expressed by P, it is not enough to attach the predicate P to them. There is something more to take into account. We must add an explanation for why those two things share P. If we analyse (1) and (2), we cannot find a common explanation for their truth; in other words, there is no explanation that can tell us, at the same time, why it is true that Rome is the capital of Italy and why it is true that snow is white. As a matter of fact, what explains the truth of (1) is the fact that Rome is the capital of Italy, and what explains that Rome is the capital of Italy is the political history of Italy. On the other hand, what explains the truth of (2) is the fact that snow is white, and what explains that snow is white is the nature of snow. Of course, the political history of Italy has nothing to do with the nature of snow. That is why the two sentences don't actually share any property.¹⁹

Therefore, what leads a deflationist to claiming that truth is not a property – or, at least, not a property in the strongest sense of the word – is exactly the lack of this latter requirement. In fact, the argument has as conclusion that truth is not a property *tout court*. However, contemporary deflationists support a less radical view, according to which truth is not a substantive property, because it doesn't have an underlying nature. As a matter of fact, we cannot find out a common explanation for why all true things are true. There is no genuine and substantive phenomenon of truth and of being true that such predicate describes, this is what deflationism maintains. In the words of Horwich:

Insofar as “true” does not have the role and meaning-constituting use of an empirical predicate, we can appreciate *a priori* that there will be no reductive analysis of truth to empirical properties. Similarly, we can see that the fundamental facts about truth will not be natural laws relating it to empirical phenomena. [...] This is what is meant by saying that the property of truth is not “substantive”. [Horwich 2008, p.31 (emphasis in original)]

In spite of that, truth still is a property of a kind. Which kind? According to some of the most important advocates of contemporary deflationism, such as Field,²⁰ Horwich²¹ and Beall,²² truth is a *logical* property. To be a logical predicate means that truth doesn't hold in virtue of the nature of a specific object, i.e. the substantive property of truth. More specifically, here to be a logical property means that truth satisfies some logical needs: it allows us to formulate certain infinite conjunctions and disjunctions that we cannot express otherwise.²³

We can effectively summarize this section with the help of Beall's words:

¹⁹ Stoljar and Damnjanovic[2014].

²⁰ Field[1994a], p.264.

²¹ Horwich[1998], p.2.

²² Beall[2009], p.12.

²³ The detailed explanation of the utility of the truth predicate is delayed to next section.

The device [i.e. truth] tags no substantive “nature”, and wasn’t intended to do so. As so-called disquotationalists have long said, “true” [...] is unlike ordinary predicates, which are introduced to “name” some feature of the world. Our device “true” was not introduced to name any feature of the world; it is simply a tool constructed to facilitate the use of our ordinary predicates and language, generally. [Beall 2009, p.2]

2.1.4 The Term Thesis: Truth as a Device of Semantic Descent

The term thesis holds that truth – i.e. expressions such as “is true” and “is false” – has a fundamental role within our language, despite the fact that, according to deflationary theories of truth, there is no substantive property associated with the truth predicate. What is this function performed by truth? According to deflationism, the role of the truth predicate in our language is merely expressive, and, more specifically, it allows us to formulate some useful generalizations, that is, infinite conjunctions and disjunctions that we cannot express in any other way because of the expressive limitedness of our language due to our finite nature. This function can be summed up in a nutshell by saying that the truth predicate is a device of semantic descent. The expression “semantic descent” grasps the idea that the truth predicate cancels out the semantic ascent that we achieve by forming the quotation name of a sentence. Thus, the semantic ascent allows us to switch from the use of a sentence to its mention; in other words, the semantic ascent allows us to switch from the things a sentence is about to the sentence itself. The truth predicate cancels out the semantic ascent by directly presenting the content of the sentences without their quotation names.

In particular, as a device of semantic descent, the truth predicate performs two very important expressive functions: it enables the formulation of generalizations that would otherwise require a device allowing the formulation of a particular kind of infinite conjunctions and disjunctions; and, in addition, it allows blind ascriptions of truth to take place, that is, it enables the indirect expression of commitment to claims towards which we can’t express that commitment in a direct way. We can restate what being a device of semantic descent means for the truth predicate using, once more, Quine’s words:

By calling the sentence [i.e. “snow is white”] true, we call snow white. The truth predicate is a device of disquotation. We may affirm the single sentence by just uttering it, unaided by quotation or by the truth-predicate; but if we want to affirm some infinite lot of sentence that we can demarcate only by talking about the sentences, then the truth predicate has its use. We need it to restore the effect of objective reference when for the sake of some generalization we have resorted to semantic ascent. [Quine 1970, p. 12]

Therefore, the truth predicate allows us to expand the expressive power of our language in two different ways. First, when we try to achieve certain generalizations that force us to talk of sentences even if the focus of our talk actually is the reality. For instance, in sentences such as “‘Snow is white’ is true” we can easily avoid the truth predicate, by simply and directly saying that snow

is white. “Snow is white’ is true” says no more nor less than what “Snow is white” says. As a matter of fact, in sentences like these the point of our talk is not the quoted sentence, but rather the fact in the world which the quoted sentence refers to. The following example, instead, is different because the shift from mention to use of a sentence does not appear equally clear and, as we will see, is neither possible to achieve in our natural language:

(3) Every sentence of the form “P or not P” is true.²⁴

How can we reword (3) without using the truth predicate? We can’t do that. We can’t avoid the use of the truth predicate by saying “Every sentence of the form ‘P or not P’”, or, alternatively, by simply saying “P or not P”, because the content we want to convey using (3) is different from that expressed by these two sentences. With an utterance of (3) we are expressing our commitment to all the claims of the form “P or not P”, hence, the appropriate way to communicate the same commitment without the aid of the truth predicate would be expressing the conjunction of all the sentences of the form “P or not P”. The problem is that such conjunction is infinite and the only way we have to express this kind of infinite conjunction within a natural language is by means of the truth predicate.²⁵

The same goes for generalizations that wish to grasp the meaning of infinite disjunctions. It is worth noting that this analysis doesn’t hold only for cases of infinite conjunctions and disjunctions, but also for those cases where, for reasons of time and space, we cannot provide the full list of claims we want to talk about.

As we just saw, there are several reasons why we might want a device that allows us to formulate infinite conjunctions and disjunctions. It is worth noting that there are other reasons pinpointed by Field to be added to those I already mentioned.²⁶ First, a device of generalization is really necessary just if we wish to formulate only true sentences or to have only true beliefs. In these situations we wish to express the infinite conjunction of all claims such as “I utter (believe) that ‘snow is white’ only if snow is white”. Since this conjunction is not available to us, we have to rely on the truth predicate in the following way: “I utter (believe) only true claims”. Another example in which the truth predicate, as a device of infinite conjunctions and disjunctions, is not dispensable is when one wishes to reject a theory about the physical world formulated with finitely many axioms, but she doesn’t know specifically which axioms to reject. Since she cannot deny the entire disjunction of axioms of the theory, she is forced to use the truth predicate and say: “Not every axiom of this theory is true”.

²⁴ Quine[1970], p.80.

²⁵ A means of encoding infinite conjunctions and disjunctions, that doesn’t require the use of the truth predicate, is substitutional quantification together with schematic sentence variables, but they are available only in formalized languages, and, hence, not in natural languages, that are what we are interested in. So, the truth predicate “is useful in so far as it functions as a surrogate for substitutional quantification.” [Armour-Garb and Beall(2005), p.6] In other words, the truth predicate perform for the natural language the role that substitutional quantification perform for formalized languages.

²⁶ Field[1994a], pp.264-265.

Chapter 2

In addition to cases of infinite conjunctions and disjunctions there is a second way according to which truth is necessary in virtue of its being a device of semantic descent. This second use of the truth predicate corresponds to blind ascriptions of truth, namely, all those cases where we wish to express commitment to a claim, or to a body of claims, that we don't entirely remember. For instance, let's assume we agree with everything John said yesterday, but assume also that we don't remember exactly what he said. How can we express our commitment towards John's claims? We can't, of course, list every claim John said because we don't remember all of them. The only available way is via the use of the truth predicate by saying something like that:

- (4) Everything John said yesterday is true.

Things are even clearer in cases where we can't in any way retrieve the claim we wish to endorse. This is the case of the following sentence:

- (5) The next thing John says will be true.

Neither (4) nor (5) can be restated without the use of the truth predicate, because, of course, in order to shift from the sentence to the things the sentence is about – i.e. to the world, the reality – we should know exactly what the sentence is about, but, neither in (4) nor in (5) we know all John's relevant claims. Hence, in our natural language we can semantically descend from the mention of a sentence to its use in no way but using the truth predicate, and both (4) and (5) are examples of this fact. This is exactly how the truth predicate behaves. This is exactly the fundamental role of the truth predicate that, according to deflationists, can be performed even without the ontological commitment towards a substantive property of truth.

Armour-Garb and Beall help me summarize this section:

Tarski asserted something to which you were not privy but that you are inclined to accept (on faith, as it were) and that you wish to assert yourself. Not knowing what he said, how can you do it? Suppose that he asserted "Natural languages are inconsistent". Given the relevant instances of (ES), together with the fact that Tarski asserted "Natural languages are inconsistent", you can indirectly assert what he did – that is, you can take on the same commitment – by asserting

- (1) What Tarski said is true.

[...] Related to the function of affording indirect assertions is that of affording endorsements. [...] With the truth predicate you can endorse what Tarski said by asserting (1). Truth-talk affords the means for endorsing and commending claims. [Armour-Garb and Beall 2005, p.4]

In conclusion, the conceptual necessity of this kind of generalizations – i.e. infinite conjunctions and disjunctions – and opaque endorsement brings with it the expressive necessity of having a truth predicate that operates as a device of semantic descent, by allowing the expressive power of our natural language to expand.

2.1.5 The Concept Thesis: Conceptual Connections

Finally, the fourth thesis belonging to deflationism's core commitment is the concept thesis. Remember its formulation I gave at the beginning of the chapter: the concept thesis holds that the concept of truth cannot be explained at a deeper and more fundamental level using other concepts which the truth predicate is connected to. This means that deflationists are not only claiming that the T-schema has a central role in a theory of the nature of truth, that is, in order to talk of a suitable theory of truth it is essential that from that theory we can infer all the instances of the T-schema – indeed, this holds also for certain deflationism's rival theories. Instead, the deflationist is claiming something stronger than this because she also maintains that the instances of the T-schema are fundamental, from both a conceptual and an explicative point of view:

The instances of the equivalence schema

(ES) It is true that p iff p (= That p is true iff p)

are fundamental. That is, T-deflationists claim that there is no deeper explanation, in terms of other concepts, for why these equivalences hold. [Armour-Garb and Woodbridge 2015, p.133]

In other words, that the instances of the T-schema are conceptually and explanatory fundamental means that it is neither necessary nor possible to provide a further analysis of the truth predicate that goes beyond such instances. This means that the instances of the T-schema are conceptually *brute*, namely, they don't follow from other concepts and explain the truth-talk in a fundamental and exhaustive way without the help of anything else.²⁷ We can restate the issue as follows. The instances of the T-schema are conceptually fundamental because the concept of truth is fully determined by our disposition to accept all instances of the T-schema, and they are explanatory fundamental because they are all we need in order to explain our uses of "true". In particular, we can add the following remarks about the conceptual fundamentality of the instances of the T-schema:

The upshot of such conceptual fundamentality is that the instances of (ES) are more or less analytic, as well as being necessary and a priori. [Armour-Garb and Beall 2005, p.3]

Therefore, all we can say about the concept of truth is limited to the T-schema – i.e. its instances – and no further and deeper analysis, which could help us in a better understanding of this concept at a more fundamental level, can be provided.

To briefly sum up the section, the truth-talk is explained in a fundamental way by the instances of the T-schema. This is important, because, as a result of this view, we have that the truth conditions of a sentence don't play anymore a substantial role in the explanation of its meaning and its content.

²⁷ Armour-Garb and Beall[2005].

2.2 Varieties of Deflationism

I have, so far, put forward the main desiderata a theory must achieve in order to be defined as a deflationary theory of truth. As I have already said, “deflationism” is a label that identifies a *genus* of theories that can be divided into a variety of different *species*. On the one hand, each of these varieties of deflationism shares with the others certain general features that make it a deflationary theory of truth, and, on the other hand, it owns a number of special characteristics that differentiate it from the other species of deflationism. This often means that an objection to a specific version of deflationism isn’t an objection to deflationism as a genus.

Let’s now have a look at some of the most known and developed varieties of deflationism: disquotationalism, minimalism and prosententialism. At the end of the section, I will very briefly mention also a pair of less known theories, namely, the redundancy and the performative theory of truth.

2.2.1 Disquotationalism

Disquotationalism (also known as *disquotational theory of truth*) is one of the most important and known deflationary theories of truth, endorsed by a number of contemporary philosophers and logicians, such as, among the others, Hartry Field²⁸ and J.C. Beall.²⁹ Historically, the reference to the American philosopher W.V. Quine is essential,³⁰ because he is almost universally considered the first developer of this particular version of deflationism.³¹ According to Quine, the quotation turns the sentence into something else, namely, the name of that sentence. What the truth predicate does is the inverse operation: it allows us to shift from the name of a sentence to the sentence itself. Therefore, we understand what means that a sentence is true in exactly the same way as we understand the sentence itself. The truth predicate, in this sense, adds nothing to the original sentence. As Quine puts it:

Ascription of truth just cancels quotation marks. This is disquotation.
[Quine 1992, p.80]

Hence, this is the reason of the name “disquotationalism” to point at the version of deflationism based on that Quinean idea. In particular, according to disquotationalism, all there is to say about the nature of truth reduces to the famous and provocative Quinean slogan: “Truth is disquotation”.³²

²⁸ Field[1994a, 1994b and 2008].

²⁹ Beall[2005 and 2009].

³⁰ Quine[1970 and 1992].

³¹ Despite the main reading of the Quinean account interprets it as deflationary, there are some interpreters who claim that recognizing the disquotational role of the truth predicate doesn’t automatically lead, as consequence, to embrace a deflationary theory of truth. For instance, Joshua F. Schwartz, in his PhD dissertation, defends the following thesis: “Quine do not think of disquotation or T-sentences as the *explanans* of a theory of truth, but, building on Tarski’s work, he treats them as the *explanandum* of a rigorous inquiry into the nature of disquotation.” [Schwartz(2014)]

³² Quine[1987, p.213 and 1992, p.80].

It's important to highlight that, from a disquotational point of view, the equivalence established by the T-schema is among sentences and, as a result, truth turns out to be a property of sentences – or, to be more precise, of sentence-tokens. The reason to take sentence-tokens as primary truth-bearers, instead of sentence-types, is that, in doing so, disquotationalism can accommodate ambiguous sentences as well as sentences incorporating indexical or demonstrative expressions, whose meaning changes depending on the context. Field is unequivocal on this point:

So far I have been a bit cavalier about whether a disquotational notion of truth applies to sentence-types or to utterances. I want it to apply to utterances, so that if an ambiguous sentence or a sentence containing an indexical or demonstrative is uttered on different occasions, we can regard some utterances of it as true and others as false. [Field 1994a, p. 278]

However, we will see in next section that the main point of disagreement between disquotational and minimalist theories of truth is exactly which objects are to be taken as primary truth-bearers. The aim of choosing sentence-tokens instead of propositions is to avoid what Quine calls “creatures of darkness”,³³ which in general correspond, according to the American philosopher, to all intensional entities, among which there are propositions. This is also the view supported by Field, as we can read in the following quote:

I think that pure disquotational truth is better for these purposes for two reasons. The first and less important is ontological: it is better to avoid postulating strange entities unnecessarily. The more important reason is that unless one is very careful to limit the use made of the notion of expressing a proposition, the introduction of such a notion of proposition would beg the question in favor of inflationism. [Field 1994a, pp.266-267]

In addition to the value of avoiding this “creatures of darkness”, the choice of a deflationary theory of truth based on sentences instead of propositions seems to have deeper reasons. It is possible to introduce a deflationary notion of what means to express a proposition, for instance by saying that all the claims of the form “p’ expresses the proposition that p” are necessary truths. According to deflationism, this is the only valid and legitimate notion of proposition, but inflationary theorists obviously disagree. Using a notion instead of another without providing further clarifications would beg the question against inflationism. Therefore, avoiding these problematic entities allows deflationary theorists to move ahead into the analysis without being forced to take into account the issue of choosing and motivating the proper notion of proposition.

As we said at the beginning of the section, several philosophers choose to embrace and develop a disquotational theory of truth. The most known version of disquotationalism is probably the one developed by Field. In this view, we can apply the truth predicate in a direct way only to sentences that we understand. In fact, in order to emphasize this important cognitive requirement

³³ Quine[1956], p.180

that allows truth to perform its linguistic functions, Field renamed the truth predicate as follows: “true-as-she-understands-it” instead of simply “true”. In other words, the truth predicate, according to Field’s view, can be applied only to interpreted utterances. In the words of Field:

A person can meaningfully apply “true” in the pure disquotational sense only to utterances that he has some understanding of; and for such an utterance *u*, the claim that *u* is true (true-as-he-understands-it) is cognitively equivalent (for the person) to *u* itself (as he understands it). [Field 1994a, p.250]

Therefore, if the truth predicate can be applied only to sentences someone understands, then, as a result, the equivalence between the sentences “*p*” and “‘*p*’ is true-as-she-understands-it” is of cognitive type. Disquotational theorists admit the existence of other truth predicates that don’t meet the two conditions established by the quote above, but they require every truth predicate different from the deflationary one to be always explained in terms of the pure disquotational truth predicate.

According to disquotationalists, the reason why we have a truth predicate into our language is to allow us to express, via certain generalizations, infinite conjunctions and disjunctions that, as we know, we can’t formulate in a direct way because of our finite nature. If we were God, we would need no truth predicate in order to express those generalizations, because we would be able to list all the infinite conjuncts or disjuncts required. Since we are not God, we can achieve the same result thanks to the only means we have, namely, the truth predicate. As Field puts the issue:

The word “true” has an important logical role: it allows us to formulate certain infinite conjunctions and disjunctions that can’t be formulated otherwise. [Field 1994a, p.264]

Field keeps going along this line and sets out an important example highlighting the utility of the truth predicate. There is a version of “realism” that maintains that there might be sentences in our language that are true and that, despite this fact, we never have reason to believe. In order to formulate this form of realism we need a notion of truth. What this requirement emphasises is that the use of the truth predicate, even in this case, is merely logical. As a matter of fact, if we had the possibility to formulate in our language the infinite conjunction of sentences we wish to express, then, we might formulate the realist view without the aid of the truth predicate.³⁴

To sum up, there are two fundamental features the disquotational truth owns. The first is that we can understand “*p* is true” only if we understand the utterance *p*. the second is that “*p* is true” and *p* are cognitively equivalent. As Field explains very well, we can put the second feature of the notion of truth in the following way:

This notion is of *use-independent* property: to call “Snow is white” disquotationally true is simply to call snow white; hence it is not to attribute it a property that it wouldn’t have had if I and other English

³⁴ Field[1994a], p.264.

speakers had used words differently. [Field 1994a, p.266 (emphasis in original)]

2.2.2 Minimalism

Minimalism (also known as *minimalist conception of truth*) is, along with disquotationalism, the most discussed theory in the contemporary big picture about deflationism. Minimalism was firstly outlined and developed by the English philosopher Paul Horwich, who provides a theory with three fundamental desiderata:³⁵

- 1) The theory must give an account of truth;
- 2) it has to take into account the utility of the truth predicate;
- 3) and it must provide an account of the nature of true, as well.

Before we examine point by point these three requirements, we must first sum up very briefly the main ideas behind the minimalist theory of truth. To do so, we use Horwich's effective words:

Unlike most other predicates, "is true" is not used to attribute to certain entities (i.e. statements, beliefs, etc.) an ordinary sort of property – a characteristic whose underlying nature will account for its relations to other ingredients of reality. Therefore, unlike most other predicates, "is true" should not be expected to participate in some deep theory of that to which it refers – a theory that goes beyond a specification of what the word means. Thus its assimilation to superficially similar expressions is misleading. The role of truth is not what it seems. In fact, the truth predicate exists solely for the sake of a certain logical need. [...] we may wish to cover infinitely many propositions (in the course of generalizing) and simply can't have all of them in mind. [Horwich 1998, p.2]

In this quote, Horwich presents all the elements listed above. Let's, thus, see in detailed what the three desiderata of minimalism require.

The first requirement is, perhaps, the most interesting because the minimalist conception differs from the other versions of deflationism exactly in the concept of truth it provides. In fact, Horwich, contrary to disquotationalists, provides a propositionalist interpretation of the T-schema. According to the minimalist philosopher, understanding the truth predicate means, on the one hand, accepting every instance of the T-schema, such as the disquotationalist claims, but, on the other hand, what we must accept are the propositional instances of the T-schema, contrary to disquotationalists that maintain that we must consider the sentence-tokens as primary truth-bearers. In other words, the main difference between minimalism and disquotationalism is that the latter takes utterances – i.e. sentence-tokens – to be the fundamental bearers of truth, whereas the former takes propositions to fulfil that role.³⁶ This means that, according to Horwich, truth is a property of propositions. Hence,

³⁵ Lindström[2001].

³⁶ Despite this important and fundamental difference, we should note that disquotationalism and minimalism share a point, namely, both take the fundamental truth-bearers to be semantic objects.

the instances of the T-schema, from a minimalist point of view, have to be interpreted as being about propositions, that is, as, for instance, the following: “The proposition that snow is white is true if and only if snow is white”. As a result, the right version of the equivalence principle, in a minimalist view, can be formalized as follows:

(EP_M) The proposition that p is true iff p.

(EP_M) is everything we can say about truth, according to minimalism. In other words, the content of the truth predicate or, alternative, of expressions such as “being true” and “being false”, is exhausted by (EP_M). In addition, we should note that Horwich takes the instances of (EP_M) as material biconditionals,³⁷ which hold of necessity. In other words, this means that the equivalence between the left and the right side of the equivalence schema is necessary.

Moreover, Horwich himself emphasises that ultimately the truth predicate always applies to propositions and not to sentences, even if not in a direct way:

The ordinary predicate is primarily applied to such thing as “what John believes”, “what we are supposing for the sake of argument”, “what Mary was attempting to express”, and so on: – that is, not to sentences, but to propositions. [Horwich 2008, p.40]

The second aim Horwich demands to its minimalist theory of truth is strictly tied to the term thesis. In the same way as according to deflationism as a *genus*, also minimalists claim that the truth predicate maintains a role and an utility, despite, as we know, its meaning is limited to the instances of (EP_M). To show this, take a look at the following famous example:

(6) The proposition that snow is white is true.

We can cancel out the truth predicate from this sentence via (EP_M). If any situation in which we use the truth predicate is limited to (6) and those similar to it, then the truth predicate will turn out to be redundant.³⁸ Unfortunately, there are cases where we can’t eliminate the truth predicate via the T-schema, and these situations are exactly the cases where the function performed by truth within our language arises. In fact, the role of the truth predicate in our language is to help us in the expression of sentences we couldn’t otherwise formulate.

Lastly, the third desideratum stated by Horwich holds that we should provide an analysis of the nature of truth if we wish to give an adequate minimalist theory of truth. Also this requirement is satisfied by Horwich’s theory. As a matter of fact, as we already know, according to Horwich, truth does not have an underlying nature and, so, we have to shift from the analyses of the nature of true to an account of the function of “true”. This is all we can say about the issue of the nature of truth.

³⁷ Remember that a conditional is material if it can be defined on the basis of other connectives in the following way: $A \rightarrow B =_{\text{def}} \neg A \vee B$.

³⁸ See section 2.2.4 for redundancy theories of truth.

While the last two desiderata of the Horwichian project do nothing but underlying that minimalism belongs to the family of deflationary theories of truth, the former requirement is what, in fact, differentiate minimalism from any other deflationary theory.

In conclusion, I wish to summarize the main features of Horwich's deflationary theory. According to minimalism, the meaning of expressions like "is true" is provided by our tendency to accept the instances of (EP_M),³⁹ and these instances constitute the axioms of a minimalist theory of truth and obey the following conditions:

- i. Every sentence *p* in (EP_M) must be substituted with a sentence-token of English;
- ii. All these tokens have the same interpretation;
- iii. Through this interpretation, they express a proposition;
- iv. The terms "that" and "proposition" have their standard meaning in English.⁴⁰

2.2.3 Prosententialism

Differently from what happens according to minimalism and disquotationalism, the issue here is no more about what the truth predicate applies to, namely sentences or propositions. Prosententialism is a particular version of deflationism that behaves in a different way from the deflationary theories we have mentioned so far. As a matter of fact, on one hand, prosententialism keeps a number of peculiar features of deflationism, but, on the other hand, it rejects other just as important characters. Let's see, now, how prosententialism works and why, despite what I have just said, it deserves the label of deflationary theory of truth.

Prosententialism (also known as *prosentential theory of truth*) is the theory according to which sentences involving the truth predicate function exactly like pronouns, except that instead of standing in for names, they stand in for sentences. From here on, the name *prosentences* will be used to call the product of truth-talk. In other words, the main idea of prosententialism can be sum up as follows:

A prosentence is a device for achieving anaphoric cross-reference to sentences uttered previously in a conversation, just as pronouns are devices for achieving anaphoric cross-reference to names uttered previously in a conversation. [Stoljar and Damnjanovic 2014]

The main advocate of prosententialism is Dorothy Grover,⁴¹ who claims that the truth predicate works in an analogous way as the pronoun "she" does in the following example:

³⁹ i.e. all instances of (EP_M) in all possible expansions of English, not only those actually expressible. The reason is that English is not expressively complete, whereas Horwich's minimalist theory aims to be.

⁴⁰ Armour-Garb and Beall[2005].

⁴¹ Grover[1992].

- (7) Susan is tired, but she still has to finish to write the chapter.

In (7), “she” is a pronoun that is anaphorically dependent on Susan. The same goes for truth-talk, except that there is a shift from sub-sentential to sentential level. Consider now an example of how prosentences work:

- (8) Susan is tired. That is true, but she can’t sleep yet because she has to finish to write the chapter.

As we have said, in this case, “is true” functions in the same way as “she” does, except that it stands in for sentences instead that for names. Similarly to what we have previously said about the pronoun “she”, the expression “that is true” is a prosentence that is anaphorically dependent on “Susan is tired”, i.e. it replaces a sentence-token from which it inherits its content.

As we have said, prosententialism is a variety of deflationism with peculiar and uncommon features. The main aspect that bring the difference between deflationism as we have presented it so far and prosententialism in lies in the logical structure that they assign to sentences like “P is true”. In fact, on the one hand, according to deflationism, “P is true” is a subject-predicate structure that tells us that P is true, namely, it assigns to P the (non-substantive) property of truth. On the other hand, according to prosententialism, first, “P is true” is not a sentence, but rather a prosentence, and, secondly, as a consequence of this idea, despite “P is true” has in effect a subject-predicate structure, it is not about P, but rather it anaphorically stands in for P. As a matter of fact, “she” *stands in* for the name “Susan” in (7) and (8), but we won’t say that “she” is *about* the name “Susan”. In the same way, “P is true” *stands in* for the sentence P, but it is not *about* the sentence P.

However, it is clear that, despite the differences, prosententialism is a deflationary theory of truth. In particular, Grover herself says that:

A principal claim of the prosentential theory of truth is that truth talk – all the truth talk we need – can be explained without appeal to any kind of analysis of the nature of truth. I claim the truth predicate helps facilitate certain kinds of discourse and that this role can be explained in terms of the concept of a “prosentence”. [Grover 1992, p.3]

Therefore, what makes prosententialism a deflationary theory of truth is that, according to the former as well as the latter, it is not necessary to provide an analysis of the nature of truth, because everything we can say about truth is nothing more than what we can say about the role the truth predicate accomplishes at linguistic level. That role, even in a prosentential theory of truth, is to facilitate the expression of certain kind of claims that won’t be expressible otherwise. In support of that, in her book Grover continues:

Many other truth theories assume that a sentence containing a truth predication, e.g. “That is true”, is about its antecedent sentence or an antecedent proposition. By contrast, the prosentential account is that “That is true” does not say anything about its antecedent sentence but says something about an extralinguistic subject. [Grover 1992, p.221]

Hence, truth, also from a prosentential point of view, is not a substantive property to be ascribed to sentences or propositions. In other words, when we use the truth predicate we are saying nothing about linguistic objects, such as sentences and propositions, but instead we are talking about the reality, i.e. the things in the world.

Lastly, a further reason according to which the prosentential theory of truth is regarded as a deflationary theory lies in the central role of the T-schema. Despite according to prosententialism the instances of the T-schema are *brute*, it takes them as conceptually and explanatory fundamental, in exactly the same way as the deflationary view does.

2.2.4 Other Varieties of Deflationism

The three varieties of deflationary theories just examined – i.e. disquotationalism, minimalism and prosententialism – are the most known and accepted in the contemporary philosophical panorama. However, they don't exhaust the existent varieties of deflationism. There are other less known theories that deserve the title of deflationary theories of truth. Here, we will set out the main features of a couple of them: the redundancy theory of truth and the prescriptive theory of truth.

The first theory that deserves the label of deflationary theory is the so-called *Redundancy theory of truth* that is commonly attributed to Frank P. Ramsey.⁴² As the name suggests, the advocates of this version of deflationism maintain that truth is a redundant concept, that is, a concept that we are perfectly able to manage without. To explain why, let's consider the following sentences:

- (9) Snow is white.
- (10) It is true that snow is white.

According to redundancy theorists, (9) and (10) are equivalent in a strong sense because they express exactly the same proposition. The difference between them is merely a stylistic matter. That means that there is no feature of the truth predicate – not even its role or function – that makes it indispensable. However, by so saying redundancy theorists are not denying the importance of the truth predicate for expressive purposes, but rather they mean to underline that this important feature doesn't make the predicate indispensable. According to redundancy theorists, the use of the truth predicate is merely conventional and there's no object in the world corresponding to it. As a result, there is a sense according to which we can say that the redundancy theory is a sort of nihilistic theory of truth. It is a radical version of deflationism as I characterized it so far. Specifically, it is useful to remember that the prosentential theory of truth arises as a variety of the redundancy theory, that is, it arises from the attempt to keeping together the dispensability of truth and, at the same time, a plausible explanation of how we can formulate certain linguistic expressions – such as infinite conjunctions and disjunctions, blind ascriptions, etc. – without

⁴² Ramsey[1940].

the truth predicate.⁴³ Another possible way to modify the redundancy theory is by taking p and “ p is true” to be synonymous no more, but rather by taking them as materially equivalent. With this idea in mind, we are, therefore, moving towards a minimalist perspective.⁴⁴ To be more specific, the change suggested by the minimalists is to associate the meaning of the truth predicate to the acceptance of the instances of the T-schema, letting truth to become a fully-fledged logical predicate.

The second deflationary theory we are taking into account here is the so-called *Performative theory of truth*, which was developed by the English philosopher Peter Frederick Strawson.⁴⁵ According to performative theory, (10) expresses neither the same proposition expressed by (9) – as Ramsey holds –, nor a proposition about the proposition expressed by (9) – as Tarski claims. Strawson simply maintain that (10) expresses no proposition at all. According to him, expressions like “is true” must be used primarily as performative expressions rather than descriptive, contrary to what the tradition requires. For this reason, (10), from a performative point of view, say nothing about (9), but rather it *endorses* (9). As a matter of fact, when we use a performative expression we are not making a statement, but rather we are performing an action. When we make an utterance of (10), we are not simply expressing that snow is white, but, in addition, we are expressing endorsement towards the statement that snow is white. Thus, to say that a sentence is true means performing the act of endorsing, accepting, that statement. The difference between the redundancy and the performative theory of truth appears now very clearly: according to the latter, truth has a role that establishes its indispensability, but it does not correspond to an expressive role (contrary to what redundancy theorists state), but rather a performative one.

2.3 Problems

As we saw in the previous sections, adopting a deflationary theory of truth brings some advantages that we cannot have by pursuing other theories of the nature of truth. The main advantage is that the utility of the truth predicate in our language can be explained in terms of the expressive role it performs – i.e. the role of device of semantic descent that, as I already stressed again and again, enables us to formulate certain infinite conjunctions and disjunctions – without being forced to go beyond the instances of the T-schema, that is, without being forced to look beyond a deflationary theory, for instance, by positing the existence of a substantive property of truth.

Despite this great advantage, deflationism faces some problems that involved the theory as a *genus* and not only a specific *specie* of deflationism. In

⁴³ It is not necessary that a prosentential theory of truth follows Ramsey in asserting that truth is not a property *tout court*, but rather it is common that the prosentential theory maintains that there is a property of truth, even if not a substantive one. When I say that the prosentential theory arises from the redundancy one I'm not referring to the ontological issue, but rather to the explanation of why the truth predicate seems to be indispensable in our language.

⁴⁴ See Horwich[2008].

⁴⁵ Strawson[1949].

next sections I'll give an overview over the main problems that a deflationary theory of truth has to accommodate.

2.3.1 Deflationism and The Generalization Problem

As I stressed several times, according to a deflationary theory of truth, the main role performed by the truth predicate in our natural language is being a device of generalizations, that is, a device that enables us to express infinite conjunctions and disjunctions by means of the generalization of their conjuncts and disjuncts. This feature of the truth predicate is fundamental because, in a natural language, we cannot in any way assert each of the relevant sentences – i.e. each of the conjuncts and disjuncts – and, so, truth is the only device we have in order to address this expressive lack.

The philosopher of Indian origin, Anil Gupta – followed by the Americans Scott Soames and James A. Woodbridge –,⁴⁶ claims that deflationary theories of truth are subject to the so-called *Generalization Problem*, which establishes that the notion of truth, in the way deflationists characterize it, fails to accomplish its main role.⁴⁷ If that is the case, the consequence is that deflationism characterizes a notion of truth that fails to respect one of the main requirements a truth predicate must respect according to the theory itself.

Let's see now the argument of the objection. Gupta, Soames and Woodbridge observe that deflationists are not able to account for the generalizations formulated using the truth predicate, contrary to what they claim. In the words of Woodbridge:

Deflationists about truth typically emphasize the logical functioning of the truth-predicate, claiming that this logical role fully explains and vindicates the notion of truth. It is interesting, then, that one of the greatest threats to deflationary views is the problem of accounting for the very generalizing role they stress as truth's central function. [Woodbridge 2003, p.285]

The reason for this difficulty is that these generalizations are not logically equivalent, in a strong sense, to the conjunctions or the disjunctions of their instances, but rather the former are logically stronger than the latter. In fact, the generalization of a conjunction (disjunction) implies the conjunction

⁴⁶ Gupta[1993], Soames[1997] and Woodbridge[2003].

⁴⁷ Gupta[1993] introduced the Generalization Problem with regard to Horwich's minimalist theory of truth, but Woodbridge[2003] showed that this problem applies also to the other major deflationary theories, i.e. disquotationalism and prosententialism. Minimalism faces the generalization problem because it uses to explain truth by appealing to set of propositional instances of the T-schema, i.e. (EP_M), that are taken as undervived axioms. The problem is that it doesn't provide a way to unify these axioms in order to create a generalization. As far as prosententialism is concerned, the generalization problem arises because quantifications, that should provide the relevant generalizations by means of the truth predicate, are substitutional, rather than objectual. The result is that they cannot be regarded as genuine generalizations, because the universal quantification always comes with a particular substitution class. Lastly, even the disquotational theory developed by Field is subject to the generalization problem for the same reason of prosententialism: quantifications of sentences, such as "For everything sayable, if John says it, then it is true" are substitutional. Hence, the generalization problem is a common problem for all the different varieties of deflationism, and, so, can be regarded as a problem for deflationism *tout court*.

(disjunction) of its instances, but the contrary is not the case. Let's consider now the sentence (3), that, remember, says:

(3) Every sentence of the form "P or not P" is true.

(3) is logically stronger than the sentence stating the conjunction of all its instances:

(11) "Snow is white or snow is not white" and "Rome is the capital of Italy or Rome is not the capital of Italy" and "Moon is made of cheese or Moon is not made of cheese", and so on.

(3) and (11) are equivalent only in a much weaker sense. They cannot be equivalent in a strong sense because of two reasons: first, they don't have the same meaning, and, second, they don't imply the same things. As a matter of fact, as I've said, universal generalizations are not logical consequences of the set of their instances, while, on the contrary, the set of their instances is logical consequence of the universal generalizations. This means that from (3) we can logically infer (11), but the contrary is not the case. Hence, a generalization and the infinite conjunction (disjunction) of its instances imply different things. In order to move from (11) to (3) we should add a requirement for (11): we must require that the quoted sentences in (11) are all sentences of the form "P or not P" and nothing else is a sentence of this sort. If we add this last claim, then we will fill the gap between (3) and (11) – and, more generally, between conjunction (disjunction) of the instances and generalization. However, this won't solve the problem because that claim is a generalization as well.⁴⁸

Of course, this is a big problem for deflationism, because the equivalence between sentences like (3) and (11), that is, between infinite conjunctions or disjunctions and their generalizations via the use of the truth predicate, is the basis of a deflationary theory of truth. Moreover, this difference at level of logical power becomes a difference at level of explanatory power: the explanation of each instance is not sufficient for the explanation of the generalization.⁴⁹ The conclusion of the analysis conducted by Gupta, Soames and Woodbridge is that we can have only a weak equivalence between instances and generalization, not an equivalence in its strong sense as the theory requires.

For the sake of clarity, we can restate the issue of the generalization problem using Soames' words:

I take this [i.e. the generalization problem] to be a genuine problem. Because the minimal theory is just a collection of instances, it is conceivable that one could know every proposition in the theory, and still be unable to infer (24) [i.e. "For any proposition p and q, the conjunction of p and q is true iff p is true and q is true], because one is ignorant about whether the propositions covered by one's instances are all the (relevant) propositions there are. For example, given only the minimal theory, one might think: perhaps there are more propositions

⁴⁸ Woodbridge[2003], p.287.

⁴⁹ Ibid. and Gupta[1993], pp. 66-67.

and maybe the predicate *true* applies differently to them. A person in such a position has no guarantee of (24), and might lack sufficient justification for accepting it. It seems doubtful that such a person should be credited with understanding the word *true*, or with grasping the notion of truth. On the contrary, it seems that grasp of the notion of truth (plus an understanding of what a conjunction is) should be sufficient for knowing (24). [Soames 1997, p.31 (emphasis in original)]

Since the grasp of the notion of truth is not sufficient in a deflationary theory of truth for knowing a generalization, such as (24), then a deflationary theory of truth does not seem a suitable theory of the nature of truth.

In conclusion, to sum up, the challenge raised by the generalization problem to deflationism is to account for general claims formulated via the truth predicate, only using the tools available to the deflationary theory. The conclusion is that deflationism can only grasp the conjunction of the instances of the T-schema, but it can't grasp also their generalization, i.e. it has no tool to derive such generalization from the instances.⁵⁰

2.3.2 Deflationism and Truth-value Gaps

One of the main problems a deflationary theory of truth has to deal with is the putative incompatibility between deflationism and truth-value gaps. More specifically, the issue can be sum up as follows: if there is no substantive property corresponding to the truth predicate, how can we account for the several cases of indeterminacy? There are, in fact, sentences whose truth-values cannot be determined for sure, because of different underlying reasons. Postulating a gap between truth and falsity traditionally solves these cases of indeterminacy. However, this is a problem for a deflationary theory of truth: how can we account for vagueness' situations that appear because of the use of vague predicates – that is, borderline cases, such as “being tall” or “being bald” – that don't have a straight line dividing truth from falsity? Or, how can we account for some moral propositions for which there isn't a sharp distinction between truth and falsity, like, for instance, when we wonder whether we should return a favour or not? Or, how can we incorporate into our theory sentences whose truth-values are indeterminate because they express future contingent propositions, such as the famous Aristotelian example “There will be a sea battle tomorrow”? Finally, certain problems resulting from semantic notions, such as the liar paradox and the truth-teller,⁵¹ have a solution based on truth-value gaps; how can we account for these cases within our deflationary theory?⁵²

Therefore, it seems there is a sort of incompatibility between deflationism and truth-value gaps that generates an important problem for a deflationist because, as we saw, truth-value gaps are usually involved in the solution of several problematic situations. As Beall puts the issue:

⁵⁰ See Armour-Garb[2004] and Armour-Garb and Beall[2005].

⁵¹ The truth-teller is a sentence that says of itself that is true. It is not a paradoxical sentence, but still is a kind of semantic pathology.

⁵² Other examples that are commonly considered paradigmatic cases of truth-value gaps are sentences involving presupposition failure and incomplete situations.

Gaps seem to be ubiquitous (vagueness, perhaps ethical discourse, comical discourse, etc.); if deflationism cannot admit gaps then this tells heavily against deflationism. [Beall 2002, p.300]

Hence, either deflationism is compatible with truth-value gaps, or an alternative way to account for indeterminacy is needed for it.

Let's now see why deflationism and truth-value gaps seem to be incompatible. The rule thesis of the deflationism's core commitment holds, remember, that the use of a truth-bearer and the use of an ascription of truth to it are intersubstitutable in all transparent contexts. As we saw, this requirement can be semantically achieved via the validity of the equivalence principle and the principle of intersubstitutivity. Recall the equivalence principle for truth:

(EP) $\langle p \rangle$ is true iff p .

Now, the advocates of gappiness maintain that there are examples of sentences that we cannot evaluate, that is, there are example of meaningful sentences that are neither true nor false. The gappiness claim can be summarized in the following way:

(G) not (either $\langle p \rangle$ is true or $\langle p \rangle$ is false)

Let's now consider the argument stating that the assertion that there are truth-value gaps leads to contradiction if we assume deflationism with all its logical requirements:

- | | | |
|----|---|---|
| 1) | $\neg(T(\langle p \rangle) \vee F(\langle p \rangle))$ | [(G) by assumption] |
| 2) | $\neg(T(\langle p \rangle) \vee \neg T(\langle p \rangle))$ | [by 1 and the equivalence $F(\langle \alpha \rangle) \equiv \neg T(\langle \alpha \rangle)$ ⁵³ |
| 3) | $\neg T(\langle p \rangle) \wedge \neg \neg T(\langle p \rangle)$ | [by 2 and De Morgan] |
| 4) | $\neg p \wedge \neg \neg p$ | [by 3 and (EP)] |
| 5) | $p \wedge \neg p$ | [by 4 and Principle of Double Negation] |

We just proved that from the assumption that there are gappy sentences, along with the principles valid in a deflationary theory of truth, it's very easy to infer a contradiction. This, of course, is a problem for the majority of logicians⁵⁴ and, according to them, it must be interpreted as a symptom of the fact that deflationism and truth-value gaps are mutually incompatible.

Traditionally, the answer to the problem faced by the majority of deflationists – including Horwich – implies to block the derivation of the contradiction at step 4 where the principle of equivalence is involved. According to Horwich, in this principle we are using the strict material conditional, namely, a conditional that holds of necessity. This means that we can only substitute a sentence with an attribution of truth to it *salva veritate* – i.e. we can substitute “ p ” with “ $\langle p \rangle$ is true”, and vice versa, only if they have the

⁵³ Remember that this equivalence holds only for deflationists. In fact, other theorists, such as Priest, reject the equivalence, and in particular reject the move from $T(\neg \alpha)$ to $\neg T(\alpha)$. We will see the details in next chapters.

⁵⁴ Except that for dialetheists, of course, who accept the existence of true contradictions by definition.

same truth-value. However, it seems that this requirement is not respected in the argument above because the starting assumption states that there might be sentences that are neither true nor false. If, in fact, we fill the T-schema – or, equivalently, the principle of equivalence – with a sentence that lack of truth-value, what we obtain is that the right side of (EP) is neither true nor false, but the left side is clearly false, because it says of a sentence that is neither true nor false that it is true. Hence, the two sides of the principle don't have the same truth-value when we are dealing with truth-value gaps and this is the reason why Horwich doesn't allow the validity of the equivalence principle for gappy sentences.

Another possible way to avoid the problem is the one developed by Beall, which requires a distinction between weak (\sim) and strong (\neg) negation.⁵⁵ If we consider the classical truth-values, i.e. “true” and “false”, we have that both kinds of negation invert the truth-values. The difference appears, as we expect, when we are dealing with gappy sentences: weakly negating a gappy sentence generates another gappy sentence, while strongly negating a gappy sentence produces a true sentence. The fundamental difference is that falsity now is defined as truth of weak negation. To say of a sentence that is gappy means to say $\neg(T(\langle p \rangle) \vee F(\langle p \rangle))$, but since now falsity is equivalent, by definition, to weak negation, what we obtain is $\neg(T(\langle p \rangle) \vee \sim T(\langle p \rangle))$. Let's try to restate the above argument on the light of this difference:

- | | |
|--|---|
| 1) $\neg(T(\langle p \rangle) \vee F(\langle p \rangle))$ | [(G) by assumption] |
| 2) $\neg(T(\langle p \rangle) \vee \sim T(\langle p \rangle))$ | [by 1 and the equivalence $F(a) \equiv \sim T(a)$] |
| 3) $\neg T(\langle p \rangle) \wedge \neg \sim T(\langle p \rangle)$ | [by 2 and De Morgan] |
| 4) $\neg p \wedge \neg \sim p$ | [by 3 and EP-T] |

There is no contradiction at all. In fact, from $\neg \sim p$ we cannot infer p and so, we don't have any contradiction of the form $p \wedge \neg p$, contrary to what we obtained with the previous argument, where we didn't have the distinction between the two forms of negation.

To sum up, truth-value gaps point to a crucial problem for deflationism, because from the assertion that there are gaps, along with the principles valid in a deflationary theory, a contradiction immediately follows. As a result, in order to avoid the incompatibility between deflationism and truth-value gaps – incompatibility that will decrease the attractiveness of the theory because truth-value gaps meet the support of a great number of philosophers and logicians in the contemporary philosophical panorama, especially because they allow to explain certain facts that will remain unanswered without them – there seem to be two moves available for a deflationist: the first is to restrict the T-schema only to sentences that allow the substitution of “ p ” and “ $\langle p \rangle$ is true” *salva veritate*; the second is to take ordinary negation to be ambiguous and, as consequence, to admit the existence of two different negations.⁵⁶ However, both answers are vulnerable to several problems.

⁵⁵ Beall[2002].

⁵⁶ There are, of course, other possible solutions available to deflationism that have been developed over the years. For instance, we can maintain that gappy sentences are ill-formed. According to that strategy, there is no problem for deflationism because these sentences simply

First, deflationism is a theory totally based on the equivalence principle and, so, restricting the T-schema means restricting the meaning of truth itself, because, as we know, deflationists claim that the meaning of truth is totally grasped by the instances of the T-schema. In other words, there's nothing more in the meaning of truth but the instances of the T-schema.

Moreover, the restriction of the T-schema leads to another kind of problem. Let's consider the following sentence:

(12) Everything the Pope said is true.

Via the generalization expressed by (12) we express our agreement with everything the Pope said, whatever it is. The problem arises when, for example, among the propositions said by the Pope there is also a moral proposition. We wish all instances of the equivalence principle to hold – also the instances for moral propositions that, according to most people, are neither true nor false – because otherwise with (12) we couldn't express everything we want to express.⁵⁷

Another problem is connected with the strategy that prescribes a difference between weak and strong negation: disambiguating negation means, in fact, changing the meaning that we commonly attribute to this connective, and by so doing we are violating the methodological requirement of classical recapture.

Lastly, even if we distinguish the two meaning of negations, the problem is far from being solved because there still is a kind of negation – i.e. strong negation – that makes deflationism and truth-value gaps incompatible.

2.3.3 Deflationism and The Liar

The main problem for a deflationary theory of truth is the well-known liar paradox. In fact, the liar questions the cornerstone of deflationism, namely, the T-schema. As we saw in section 1.4.2, this paradox points to a crucial problem for all of those theories that wish to keep the T-schema safe and possibly unrestricted.

Let's recall the argument and consider the following formalization of the T-schema:

(T) $T(\langle a \rangle) \leftrightarrow a$

don't express any proposition. Another possibility is to claim that gappy sentences don't exist *tout court*. The phenomenon that made us believe that certain sentences lack of truth-value must actually be explained in another way. Another possible strategy is to abandon classical logic in favour of a non-classical one. We have two possibilities to achieve this result: on the one hand, we can restrict the law of excluded middle and, consequently, we don't accept its instances in the case of gappy sentences – this is the case of paracomplete logics –; on the other hand, we can reject the *ex contradiction quodlibet* and, in this way, we can accept the inconsistency of our language and, as a result, we can simply accept the contradiction due to the union of deflationism and truth-value gaps – this is the case of paraconsistent logics. For a general overview of all these and other strategies, see Stoljar and Damnjanovic[2014].

⁵⁷ See Stoljar and Damnjanovic[2014].

Let's now consider the instance of the T-schema for the liar that, remember, can be formalized as follows:

$$(L) \quad F(\langle L \rangle)$$

and we obtain:

$$T(\langle L \rangle) \leftrightarrow L$$

by what (L) says, i.e. the equivalence between L and $F(\langle L \rangle)$, this is equivalent to:

$$T(\langle L \rangle) \leftrightarrow F(\langle L \rangle)$$

From this implicit contradiction we can easily derive an explicit one of the form:

$$T(\langle L \rangle) \wedge F(\langle L \rangle)$$

Therefore, it is now clear why the liar poses a serious problem to deflationism: it is incompatible with one of its key pillar, namely, the T-schema. As we saw, in fact, from the T-schema and the liar a contradiction immediately follows.

One of the reasons why the liar represents a particularly problematic case for a deflationary theory of truth is that the paradoxality of the liar cannot be cancel out in the traditional way, by taking the liar as lacking of truth-value, that is, by taking the liar as neither true nor false. The reason is that truth-value gaps, in turn, pose a serious problem for deflationism, as we saw in the previous section. As a matter of fact, if we assert that the liar is neither true nor false, then we are only displacing the problem, because, as we saw, this move doesn't avoid the contradiction.⁵⁸ Moreover, there is another reason why the appeal to truth-value gaps doesn't solve the problem. In fact, this solution is subject to the revenge phenomenon that produces a strengthened version of the liar, i.e. the sentence that says of itself that it is not true, or the sentence that says of itself that it is false or neither true nor false. In both cases, the appeal to truth-value gaps is of no help in the resolution of the paradox because we can't avoid or solve the paradoxality of these two revenge liars by saying that they are neither true nor false.

An alternative way to account for the status and the problem raised by semantic paradoxes is to take them as meaningless, that is, to claim that semantic paradoxes express no proposition at all.⁵⁹ This is known with the

⁵⁸ Things would be different if we could restrict the T-schema. However, as we stressed again and again, the full T-schema is an unavoidable desideratum for a deflationary theory of truth. Hence, the way of restricting it is not desirable for a deflationist.

⁵⁹ Actually, among the meaningless sentences there is also the truth-teller, that is, the sentence that says of itself that is true. The truth-teller, of course, doesn't share the paradoxical character of the liar, but there is a sense according to which it seems to be ill-formed in a similar way. In the case of the truth-teller we can arbitrarily choose which truth-value to give to the sentence. If we assume that the truth-teller is true, then this is exactly what the sentence says and, hence, it turns out to be true. If, on the contrary, we assume that the truth-teller is false, then for what the sentence says – i.e. the truth-teller says of itself that it is true –, it turns out to be false. There is no paradoxality in the behaviour of the truth-teller, but there still is something awkward in it. In

name of *meaningless strategy* and is defended by J.C. Beall.⁶⁰ As Armour-Garb puts it, this strategy seems to be intuitively satisfactorily because:

It captures many people's initial response to the paradoxes; and it is theoretically important: if successful, it might allow the deflationist to avoid the revenge problem that plagues so many putative solutions to the Liar. [Armour-Garb 2001, p.280]

Therefore, in addition to solving the liar, the meaningless strategy allows us to provide a solution also to its strengthened version, i.e. the revenge liar. However, there is a serious problem that the meaningless strategy must face. The solution to the liar that its advocates provide doesn't seem to be plausible at all. The reason is that it is in virtue of the meaning of these defective sentences – i.e. semantic paradoxes and other pathological sentences – that we recognize their pathological or paradoxical character and behaviour.⁶¹ If these sentences were meaningless, we wouldn't understand what they say and, hence, we couldn't infer the contradiction or their pathological character, that is what enables us to point to them as paradoxes or ill-formed, and, so, as problematic for our theory.⁶²

Horwich decides to adopt another strategy in order to provide an adequate solution to the problem raised by the liar. He proposes to restrict the class of correct instances of the equivalence principle. Hence, on the one hand, Horwich takes the liar as meaningful – that is, he admits the liar to express a proposition –, but, on the other hand, he proposes to exclude it and the other problematic propositions from the list of available instances of the equivalence principle.

Horwich's solution faces, straight away, a methodological difficulty: how can we decide which are the acceptable instances of the T-schema and which are not? Horwich himself provides the conditions a theory must meet in order to adopt his strategy. These requirements can be summarized in the following three points:

- i. The theory doesn't have to generate contradictions, such as the one obtained from the liar sentence;
- ii. The list of excluded instances of the T-schema must be as short as possible;
- iii. The theory must provide an explanation of why these instances have to be unacceptable, i.e. the move of excluding some instances instead of others must not be *ad hoc*.⁶³

With these desiderata in mind, we should provide a theory that solves the problem of the semantic paradoxes and that, at the same time, answers the previous question. Unfortunately, as we will now see, the problem is far from being solved, because those requirements only shift the problem from one

fact, the truth-teller is somehow ill-formed, because whichever truth-value we decide it must have, then it will always be fulfilled by the sentence.

⁶⁰ Beall[2001].

⁶¹ Or, we can say, their ill-formed character, as in the case of the truth-teller.

⁶² Armour-Garb[2001], pp.283-284.

⁶³ Horwich[1998], p.42.

place to the other. In fact, now, the problem becomes to provide an answer to the first of the previous criteria: which are the unacceptable instances?

Let's try to figure out what the answer to this question might be. Perhaps, the intuitive solution is to accept all non-paradoxical instances of the T-schema. In this case the first condition is satisfied because it is explicitly stated that the unacceptable instances are paradoxical sentences, or propositions, as the theory prefers. What about the last two requirements? They do not seem to be met. As far as the second criterion is concerned, the list of excluded instances is, in fact, short, but the problem is that the list is actually too much short, because it doesn't include the non-paradoxical but still problematic cases, such as the truth-teller. Similarly, also the latter desideratum is not satisfied, because an independent explanation of why the paradoxical instances have to be excluded from the T-schema cannot be provided. Therefore, the only reason seems to be their paradoxality, but this makes the decision *ad hoc* and, for this reason, unacceptable.

This is not the only problem faced by Horwich's proposal. There is a further issue that must be addressed, and it is well explained by Armour-Garb and Beall:

Which sentences are the non-paradoxical ones? Unfortunately, Horwich does not tell us. The reason Horwich does not tell us is that there is no known answer to the question. [Armour-Garb and Beall 2003a, p.383]

It is impossible to a priori know which sentences are paradoxical and which are not. Of course, there are well-known paradoxical sentences, such as the liar and its kin, but there are also many other paradoxical sentences of which we can't provide a list and of which we can't provide the conditions for paradoxality.⁶⁴

For all these reasons, Horwich's proposal seems to be very hard to achieve. Even if his solution were modified in order to exclude not only paradoxical but all ungrounded propositions,⁶⁵ it would, however, face a number of problems. In fact, there are independent reasons why excluding some instance of the T-schema is not a suitable solution. First, there seems to be some competent speaker who, on the one hand, endorses even those controversial instances and, on the other hand, is still able to account for the

⁶⁴ After all, Kripke[1975] showed that there isn't an intrinsic criterion that will enable us to identify paradoxical sentences. This means that paradoxes arise even when the form of the sentences is not manifestly paradoxical. An example to support this idea is that many of our ordinary assertions involving truth and falsity turn out to be paradoxical if the empirical facts are extremely unfavourable. This fact, according to Kripke, proves that there isn't a syntactic or semantic criterion that enables us to distinguish between "bad" and "good" cases. If Horwich's strategy to solve the problem of semantic paradoxes is founded on the restriction of the instances of the T-schema to non-paradoxical sentences, then the lack of a criterion of this sort represents a real problem for it.

⁶⁵ *Groundness* is a property of sentences, according to which a sentence is grounded if and only if either it refers to non-truth-theoretic facts or it refers to other sentences that are grounded. Otherwise, the sentence is *ungrounded*. This vocabulary is due to Kripke[1975], who provides a particular construction – i.e. the construction of the fixed point – in order to determine which sentences are grounded and which are not. The introduction of this property will help Horwich to solve the problem its strategy faces of the ill-formed sentences that are not paradoxical, such as the truth-teller. The truth-teller is, in fact, clearly ungrounded because it refers only to itself, so, it refers to nothing but truth-theoretic facts. In this way, the truth-teller is one of the propositions which the T-schema cannot be applied to.

problem these instances raise. This is the case of dialetheism that, as we know, admits the existence of true contradictions and, hence, is perfectly compatible with the view that we should unrestrictedly accept the equivalence principle. Secondly, by restricting the T-schema, it becomes more difficult explaining why these sentences are pathological, because there really is something in the concept of truth that explains why these paradoxes arise. However, according to deflationism, the concept of truth is a thin concept, namely, a concept without an underlying nature, and its meaning is fully explained by the T-schema. Hence, if the instance of the T-schema for the liar doesn't hold, the concept of truth won't be able to give us an explanation for its paradoxicality, i.e. the T-schema won't explain how and why the liar differs from the unproblematic sentences. As a matter of fact, it seems that the Tarskian biconditional for the liar grasps something right and important about the liar itself. If we restrict the instances of the T-schema, we will lose this ability. Hence, as Gupta points out:

The T-biconditionals for paradoxical sentences play an indispensable role in ordinary uses of some unproblematic sentences. [Gupta 2005, p.137]

To show what this indispensable role of the T-schema is, let's consider, for instance, the following example:

(13) For all sentences x , if x is true then " $x \ \& \ x$ " is true.

The T-schema allows us to use (13) in order to generalize the following sentence:

(14) If A , then $A \ \& \ A$.

The point here is that (13) won't be a generalization of (14) if we restrict the valid instances of the T-schema to non-paradoxical or ungrounded sentences.⁶⁶ The reason is that without the instance of the T-schema for the liar we can't shift from the sentence "If 'The liar is false' is true, then 'The liar is false & the liar is false' is true" – that is an instance of (13) – to the sentence "If the Liar is false, then the liar is false & the liar is false" – that is an instance of (14) –, and vice versa.⁶⁷

A third possible solution available to deflationism in order to solve the problem of the liar is denying, in the wake of Tarski, the truth predicate to be ascribed to sentences and propositions that themselves involve the truth-theoretic part of the language. However, as we can guess, this proposal faces the same problems as the previous one, because even this strategy involves the restriction of the T-schema to a class of sentences. As we have already said, the full validity of the T-schema seems to be an important – and, perhaps, a fundamental – requirement in order to keep safe the initial ideas and spirit of a deflationary theory of truth.

⁶⁶ Actually, it would be a generalization, but it wouldn't capture also non-problematic sentences.

⁶⁷ Gupta[2005], pp.136-137.

Deflationism about Truth

To sum up, therefore, the possibilities we have analysed so far that are available to deflationism in order to avoid or to solve the problem due to the liar paradox are three, each of which has some weak points: the first strategy requires to deny that sentences such as the liar actually express propositions; the second strategy depend upon the restriction of the T-schema to unproblematic sentences; and the third strategy corresponds to denying that the truth predicate can be consistently applied to sentences or propositions involving themselves that predicate, such as the liar do. There is another possible solution available to deflationism. It involves dropping out classical logic in favour of a paraconsistent or a paracomplete theory. If this strategy succeeds and which among non-classical logics is to be preferred is the argument of chapter 4. Here I've simply argued that a genuine problem for a deflationary theory arises from the liar paradox, and is nothing but of easy solution.

Capitolo 3

Truth-Theoretic Fictionalism

3.1 Fictionalist Theories

3.1.1 Introduction to Fictionalism

Fictionalism arises from the intuition that certain sentences, if taken at face value, are not true, but, at the same time, there is a sense according to which we are inclined to say that they are true. To introduce the issue take a look at the imaginary dialogue between Quine and one of his student proposed by Berto and Plebani:

Like all respectable fathers, Quine may have told his sons some story from time to time. He may have said things like: Once upon a time, there was a magic kingdom. In such a kingdom there lived a beautiful princess named Cinderella. Suppose one of his students caught him while pronouncing these words and told him: “Wait a minute, professor Quine! You just said ‘There was a magic kingdom’: how should I translate this into canonical notation? That looks like quantification over magic kingdoms, committing you to the existence of pretty mysterious entities. Didn’t you say somewhere (Quine 1948) that you prefer desert landscapes to Meinongian jungles?” [Berto and Plebani 2015, p.84]

How can we legitimately e truly say there is a magic kingdom? Obviously, the sentence “there was a magic kingdom” is not true if taken at face value, because there is no magic kingdom at all. However, it is equally obvious that what Quine is doing in the imaginary dialogue of the quote is not saying how things are in the reality, but rather to tell a story, and for this reason, it seems perfectly reasonable for him to say “there was a magic kingdom” without being committed to its literal content. In other words, there seems to be a sense according to which “there was a magic kingdom” is true. More specifically, we can say that that sentence is literally false but is true inside Cinderella’s story.

Fictionalism is the metaphysical view that utterances and sentences belonging to a specific area of discourse must not be taken at face value, but rather as part of a fiction. In other words, the sentences of that area of discourse are not true if taken literally, but are only fictionalistically true. As Sainsbury puts it:

Fictionalists say that some thoughts or saying are, or can be, or should be regarded as a fiction: the thoughts have value and importance, but, as with fiction in the ordinary sense, this does not consist in their being true. [Sainsbury 2009, p.152]

Hence, the first important feature of fictionalism that is worth noting is that it is an approach that applies to discourse. While dialetheism of chapter 1 is a metaphysical theory of which I'm most interested in its logico-semantic aspects, and deflationism can be meant in a broader sense as an ontological theory, fictionalism belong to a different category of theories because when one endorses a fictionalist approach she is, in fact, theorizing about a specific region of the discourse.

A second very important feature of fictionalism is that fictionalist approaches, as we have already suspected, suggest an analogy between fictional discourse, i.e. the discourse representative of the works of fiction, and some problematic areas of discourse that involve names referring to problematic entities which we don't want to be ontologically committed to. For instance, fictionalists about mathematical discourse maintain that sentences such as " π is irrational" are semantically similar to sentences belonging to fiction, such as, for instance, "Sherlock Holmes is a detective". Both sentences, if taken at face value, are false, but as the latter is true according to Doyle's stories of Sherlock Holmes, in a similar way the former is true according to the story of standard mathematics. Despite this common ground, there is a crucial difference between the two talks: mathematical discourse seems to be indispensable for our scientific theories, contrary to discourse about fictions, that seems to be perfectly dispensable. For this reason, the main problem for a fictionalist is to explain the applicability of mathematics, that is, to explain how it is possible to use mathematics despite its sentences are literally and plainly false.¹

The third feature of fictionalism that is worth mentioning is that according to fictionalist approaches there is a difference between the ontological commitments conveyed by the semantic content of a sentence and the ontological commitments of the speaker who makes an utterance of that sentence. In the former case, the ontological commitment is towards the literal content of the sentence, but, as we saw, in certain areas of the discourse what really matters is not the literal content of the sentence, but rather its fictional content, namely, its *real content* that is the real subject matter of the sentence and correspond to the content the speaker really means to convey.² Yablo explains very well the difference between literal and real content of a sentence:

What is the literal content of "the number of sheep is three times the number of goats"? Roughly this: the sheep have associated with them a number that stands to the number associated with the goats in a certain numerical relation, a relation that the number of F's bears to the number of G's only if there are three times as many F as G's. The real content is that there are three times as many sheep as goats. [...] The real content is that portion of the literal content that concerns the sheep and goats.
[Yablo 2001, p.97]

¹ As Colyvan[2011] emphasizes, according to fictionalism about mathematics, a large part of mathematics is false. This means that not all mathematical sentences are false. In fact, this is the case of negative existentials in mathematics. When we say "there is no larger prime", this sentence is not true only fictionalistically – i.e. true according to the story of mathematics –, but is also literally true because, according to fictionalism, there are no numbers, hence, *a fortiori*, there is no such thing as the largest prime number. This means that not all mathematical sentences are literally false.

² For the details about the distinction between literal and real content of a sentence see Yablo[2001].

Lastly, another important feature of a fictional perspective that is worth noting is that in most cases fictionalism is motivated by ontological scruples. What moves philosophers to embrace a fictionalist perspective of a specific area of discourse is the ontological economy. This can be meant in two different ways, as Achille Varzi points out.³ The first way is the less convincing and is called by Varzi *the Pascalian fictionalism* or *the skeptic's fictionalism*. Using an example from Varzi, even if we discovered that we are brains in a vat and that everything in the world is just an illusion, i.e. a fiction, “what goes on in that fiction matters a lot to us. Indeed that’s all that matters.”⁴ In fact, our entire life would depend from this illusion, and so we would keep pretending that it is not an illusion at all, that is, we would keep pretending that the world as we know it really exists, i.e. that cats and dogs (and all other things in the world) really exist. As Varzi puts it:

We step into a state of pretense and we feel joy or sorrow depending on how things go *in the pretense*, for what goes on in that pretense is what matters. [Varzi 2013, p.134 (emphasis in original)]

The second way to interpret the fact that ontological economy is what motivates a fictionalist perspective is the most interesting and is called by Varzi *the Berkeleyan fictionalism*. In this case, the rejection of the existence of some particular entities (e.g. universals, numbers, propositions, moral virtues, and so on) and, at the same time, the need to keep using the region of discourse involving them, is what move a philosopher in embracing a fictionalist perspective of that specific area of discourse. This second way to interpret fictionalism is most interesting because doesn't require the speaker to clearly know what there is and there is not in the world. We can keep using the relevant fragment of discourse even if the entities which it refers to don't exist, and the reason lies in the expressive utility of that region of discourse.⁵

3.1.2 First Fictionalists

Fictionalism, understood in the modern sense of the word, is a view that have spread since 1980, when the independent works of Hartry Field – about mathematical fictionalism –⁶ and Bas van Fraassen – on fictionalism about scientific theories –⁷were published. Both philosophers independently had the intuition that the aim of inquiry is not necessarily truth and when we accept a mathematical or scientific theory, we must not necessarily believe in its literal

³ Varzi[2013].

⁴ Ibid., p.134.

⁵ Actually, Varzi identifies a third variety of ontological fictionalism that, however, is the less interesting to us. I'm talking about the so-called *Humean fictionalism* that is based on the ideas of the Italian psychologist Paolo Bozzi. To be fictional in this account is the structure of reality, and not its content, like in the other two varieties of fictionalism. As Varzi explains very well:

Our propensity to give way to the unifying act of our intellect, of our imagination, makes us speak as though there were a unity, an identity, a causal nexus when in fact all we have is patterns of broken appearances. [Varzi 2013, p.141]

⁶ Field[1980].

⁷ Van Fraassen[1980].

content in order to use it.⁸ Field maintains that abstract entities don't exist, where for abstract entities we mean all those objects that lack a spatiotemporal address and have no causal powers. Hence, abstract entities are non-spatiotemporal and no-mental objects.⁹ Since mathematical objects are abstract entities, then, according to Field, we don't have to commit ourselves to their existence. In other words, Field advocates a nominalistic view of mathematical objects, namely, that mathematical objects don't exist. From this thesis it follows that all mathematical sentences are false because they involve names that lack of reference. More in details, Field's argument is based on two fundamental premises:

- i. If mathematical sentences are true when taken at face value, then mathematical entities exist and these entities are abstract.
- ii. Mathematical abstract entities don't exist (mathematical nominalism).

Hence, using the classical inference rule known as *modus tollens*, if we negate the consequent of a conditional, then we can infer the negation of the antecedent. Therefore, from premises i. and ii., by *modus tollens* it follows that all mathematical sentences are not true if taken at face value.

Field's fictionalist thesis is that mathematics, albeit it is literally and plainly false, must not be abandoned, because it has other more pregnant qualities besides truth. As Field puts it:

On this account truth isn't required for goodness; what is required instead is something called *conservativeness*, which embodies some of the features of necessary truth without involving truth. [Field 1989, p.4 (emphasis in original)]

The goodness and the value of a theory cannot be based only on its being true, but rather there are other important virtues on which to base our judgment. Hence, according to Field, mathematics must not be abandoned because its utility doesn't lie in its being true, but rather in the property of *conservativeness*. Mathematics is conservative because it allows inferring the same consequences about concreta that we would infer even without it but that would require a

⁸ Not everyone agree with the idea that Van Fraassen's and Field's theories are proper examples of fictionalism. Armour-Garb and Woodbridge underline that Van Fraassen makes no appeal to the notion of fiction and he doesn't even name fictionalism. This does not mean that we cannot construct a fictionalist theory of scientific discourse about unobservables from Van Fraassen's theory, but rather that it is not completely correct to name Van Fraassen's theory with the label "fictionalism". Similar remarks can be made for Field's nominalistic theory. For the details see Armour-Garb and Woodbridge[2015a], pp.1-2.

⁹ As Berto and Plebani[2015] well explain, abstract entities are always defined in a negative way, that is, by saying what features they lack in order to be concrete. While concrete objects can be touched because they are spatiotemporally located, and can make things happen, since they can enter into causal relation with other concrete objects, this is not the case for abstract entities. As Berto and Plebani put it:

Nobody has ever seen and touched the number 17. [Berto and Plebani 2015, p.124]

Moreover, unlike abstract objects, you can know several things about concrete objects by simply touching them and looking at them.

longer and more complex process.¹⁰ We can restate this concept by using Yablo's words:

The utility of mathematics lies in the *no-risk deductive assistance that it provides to the beleaguered theorist*. [Yablo 2005, p.91 (emphasis in original)]

Thanks to conservativeness, mathematics, thus, is “an auxiliary device that aids us in drawing inferences.”¹¹

Van Fraassen's thesis, instead, is that one might be agnostic about the claims on unobservables made by our scientific theories and that this doesn't cause any problem for the theories. In the words of Van Fraassen himself:

Such events and experiences, and such entities as sense-data, when they are not already understood in the framework of observable phenomena ordinarily recognized, are theoretical entities. They are, what is worse, the theoretical entities of an armchair psychology that cannot even rightfully claim to be scientific. I wish merely to be agnostic about the existence of the unobservable aspects of the world described by science – but sense-data, I am sure, do not exist. [Van Fraassen 1980, p.72]

Unobservables, thus, might exist or not, and, on the light of this consideration, our scientific theories that make claims about unobservables to explain certain phenomena might be literally false. Despite that, Van Fraassen maintains that we can keep refer to unobservables because the aim of science is not truth about unobservables, but rather the *empirical adequacy*, or, in other words, the representation of observables regularities among phenomena. This is why certain false or probably false scientific theories are still useful according to Van Fraassen, namely, because they are empirically adequate representations of the observable world.

As we already noticed, the main reason why philosophers might want to embrace a fictionalist perspective of a specific area of discourse is avoiding an inconvenient ontological commitment towards the entities to which that region of discourse refer. This is confirmed by the examples just stated. In fact, Field assumes a fictionalist approach of mathematical discourse because of his nominalistic view of mathematical objects, that is, because, according to him, mathematical objects, in so far as they are abstract objects, don't exist. Similar remarks can be made about Van Fraassen's view, according to which the existence of unobservable entities is questioned because we, in fact, don't know whether they exist or not and, for this reason, he decides to assume a fictionalist perspective of scientific discourse because the vocabulary of scientific discourse involves reference to objects that don't or might not exist, i.e. unobservables.¹²

¹⁰ To be more technical about the conservativeness property, we say that a mathematical theory T is *conservative* with respect to a nominalistic theory N, if every nominalistic sentence that is consequence of T+N is also consequence of N alone.

¹¹ Field[1980], p.41.

¹² The ontological reasons motivate also other varieties of fictionalism: moral fictionalism is often motivated from the fact that the natural world does not contain those entities that are moral values; similarly, fictionalism about possible worlds is often motivated from the fact that possible worlds don't exist and for this reason proposes a way to interpret the talk about them; and so on. For a brief overview on these other varieties of fictionalism see Sainsbury[2009].

To sum up, the first appearance of fictionalism may be traced back to the independent inquiries made by Field and Van Fraassen about mathematical and scientific discourses. However, over time, the scope of fictionalism has grown and nowadays it covers several other regions of discourse, such as moral-talk, modal-talk, propositional-talk, discourse about possible worlds and, among the others, very recently it has been developed to cover also the discourse about truth. This latter domain is the real subject matter of our discussion in this chapter and in the latter, but in order to talk about truth-theoretic fictionalism we first need to explain what fictionalism in general is. In the following sections I'm going to expose the main features of fictionalism, what the differences between the different kind of fictionalism are and, in particular, I aim to underline what are the philosophical and non-philosophical reasons for embracing a fictionalist perspective about a specific area of discourse.

3.1.3 Good Without Being True

As I stressed again and again, the main idea of both fictionalism's founders there is that the truth of the relevant discourse is not fundamental. If truth were fundamental for an area of discourse, we would reject that discourse on the basis that it is false.¹³ Instead, according to both Field and Van Fraassen, we can keep accepting and using the sentences of the relevant area of discourse, albeit they are plainly and literally false. As Kalderon well explains:

The aim of inquiry need not be the true representation of a putative domain of fact and that the acceptance of a theory need not involve belief in its content. [Kalderon 2005, p.2]

Therefore, truth is not the most important feature on which to base the decision over the use of a specific area of discourse. We can freely keep using these literally false sentences belonging to a region of discourse that involves words referring to non-existent objects – such as, for instance, mathematical objects according to Field and unobservables according to Van Fraassen – by relying on certain features of that region of discourse that establish its utility and that go beyond truth. Contrary to traditional belief,¹⁴ according to fictionalism it's not necessary for a theory to be true, or that the entities, which the theory theorizes about, exist, in order to enjoy the benefits of that theory.¹⁵ As Field puts it:

Truth isn't required for goodness. [Field 1989, p.4]

The above quote is often paraphrased using the most incisive motto “Good without being true.” This means that a theory can be good, or, in other words,

¹³ See footnote 1.

¹⁴ Namely, that we can't consistently believe a theory and its sentences without believing in the entities they refer to.

¹⁵ As Yablo[2005] notices, fictionalist approaches and perspectives are opposite to the one that traditionally other mathematical and philosophical theories embrace. Most of the latter, in fact, focus their analysis on what makes mathematics true, whereas fictionalism analyses what makes mathematics useful for empirical sciences.

still has a value, even if it turns out to be false. There are several features establishing the value of a theory and truth is only one among them.

According to Field, the utility of mathematical discourse lies in its being conservative with respect to the nominalistic consequences we can draw with or without it. According to Van Fraassen, the proper measure in the evaluation of a scientific theory is not its adherence to truth, but rather its empirical adequacy, whatever the use of words referring to objects that might not exist, such as the unobservables, we make.

We can use Berto and Plebani's words to briefly sum up which is the utility of a fictionalist perspective of certain regions of discourse:

This makes of it [i.e. fictionalism] an attractive strategy for those who care about ontological economy, but do not want to reduce their expressive powers. [Berto and Plebani 2015, p.87]

Fictionalism, hence, allows the ones who decide to adopt it to take at the same time both the advantages of a narrowest possible ontology and the benefits of keep using the relevant areas of discourse that involves words referring to objects not belonging to our ontology.

3.1.4 Fictionalist Attitude

One of the main features of a fictionalist approach, that is, one of the characteristic that makes fictionalism different from other perspectives about discourse is the mental attitude a fictionalist has, or must have, towards the relevant discourse, namely, her mental attitude towards that area of discourse she believe to be not literally true, but only fictionalistically true.

It's clear that the attitude towards that discourse cannot be of belief. After all, we can believe only what we believe to be true, but if we know that the sentences made into a specific area of discourse are literally false, then the attitude towards them cannot be of belief. However, there must be an attitude towards those sentences allowing us to keep using them despite they are false, that is, allowing those sentences to be of any importance for our discourse despite their falsity. What's this attitude? As the name "fictionalism" seems to suggest, this attitude is the same we have with respect to sentences made inside a work of fiction. In other words, our mental attitude towards sentences that are only fictionalistically true, such as "the Martian moons are two" for mathematical fictionalism, is the same we have towards sentences belonging to fiction, such as "Sherlock Holmes lives in London in Baker Street 221B."

Some varieties of fictionalist theories pinpoint this attitude in what is known as *make-believe*. This is the case of the famous theory of the make-believe developed by Kendall Walton,¹⁶ later used by Stephen Yablo in order to develop his own fictionalist theory.¹⁷ As Yablo explicitly puts the issue:

The fictionalist holds that we "make as if" we are asserting that S and/or believing that S and/or receiving the news that S. [Yablo 2001, p.74]

¹⁶ Walton[1990 and 1993].

¹⁷ Yablo[2001, 2002 and 2005].

Therefore, what we are doing when we utter a fictionalist sentence is merely pretending to assert and believe the literal content of that sentence. In this sense, the utterances made within the relevant area of discourse must be understood as metaphors, because we are “making as if” they are true, that is exactly the same mechanism in use in metaphors. If, in fact, we think at the assertions made within fictional works, we see that, according to fictionalists, the attitude towards them is the same we have towards assertions made within the relevant areas of discourse. When we read, in Doyle’s story, that Sherlock Holmes lives in London in Baker Street 221b, we don’t believe this to be true, but rather we simply make as if it is true, that is, we pretend that we believe it to be true. After all, Doyle’s aim is not to say the truth – that is, to report the truth about, for instance, certain historical facts that really happened –, but rather he simply wishes to tell a story, which needs not to be true and, quite the opposite, is never true by definition. We simply pretend to believe Sherlock Holmes to live in London in Baker Street 221b for the sake of the story, but we don’t believe it is really the case. The same goes for sentences belonging to relevant regions of discourse, such as “ $2+2=4$ ” for mathematical fictionalism. According to mathematical fictionalism, we don’t really believe that “ $2+2=4$ ” is true, that is, our attitude towards this sentence is not of believe, since we cannot believe something that is false and “ $2+2=4$ ” is literally false, as we saw. On the contrary, our attitude towards that sentence is of *make-believe* because we only pretend “ $2+2=4$ ” to be true and we do so because mathematical discourse is a useful device we cannot do without, or, in other words, is an aid that simplifies things about certain sentences that would be longer and more complex without it, and, in certain situations, would be impossible to express at all.

Beyond the Waltonian view of make-believe, there are other views about the right mental attitude a fictionalist has towards sentences of the relevant region of discourse. One that is worth noting is the approach pinpointing this attitude in *acceptance*.¹⁸ This thesis can be traced back to Van Fraassen,¹⁹ in particular to his discussion about scientific theories, and can be summarized in the following way with the aid of Sainsbury’s words:

The appropriate attitude is one of *acceptance*, a state that may guide action without amounting to belief. [Sainsbury 2009, p.152 (emphasis in original)]

or, alternatively, using Eklund’s words, we can say:

Van Fraassen [...] emphasizes that our attitude toward our best theory of the world is, or should be, “acceptance” rather than belief, where acceptance is an attitude that falls short of belief. [Eklund 2015]

¹⁸ Beyond the ones already cited, there are other views about the attitude towards fictional discourse. In Yablo[2006] we can find the theory of *presupposition*, according to which when we utter a mathematical sentence we are presupposing that mathematical entities exist and that, for this reason, mathematical discourse is true. In Eklund[2005] there is the theorizing that the right attitude towards fictional discourse is the *indifference*, namely, that we are indifferent to the implications of the fictional sentences about what exists and what does not exist.

¹⁹ Van Fraassen[1980].

Therefore, with an utterance of a fictionalist sentence we are not expressing belief towards that sentence, because otherwise we would believe something that we have no reason to believe. On the contrary, we merely accept it, namely, we accept that that sentence, although it is or might be not literally true, can be used to express a concept that we cannot express otherwise, or that would be more complex to express without the use of that region of discourse. In other words, a domain of discourse is associated to every area of inquiry, that is, a region of discourse that includes a series of sentences involving a specific vocabulary for that domain of inquiry. We accept a sentence of the relevant region of discourse when, as Kalderon puts it, we “have no reason to inquire further.”²⁰ Fictionalism is the view that the acceptance of those sentences belonging to the relevant area of discourse associated to the domain of inquiry of interest doesn’t imply the truth of the sentences and, for this reason, doesn’t even imply the mental attitude of belief of an agent in the content expressed by those sentences.

3.1.5 What Fictionalism Is Not

So far, we saw and we focused on what a fictionalist view about a specific region of discourse is. Let’s now see what fictionalism is not, namely, let’s now see from which other perspectives about discourse it differs. The aim of this analysis is to better understand fictionalism and to highlight its main features.

The important distinction, made in section 3.1.4, among different mental attitudes allows us to add a further aspect to the analysis of what fictionalism is or, even better, allows us to underline what fictionalism is not. In particular, the distinction between the attitude of believe and the attitude of mere acceptance, or make-believe, allows us to distinguish fictionalism from eliminativism – or, reductionism – about an area of discourse. *Eliminativism* is the view that a specific region of discourse is false and therefore must be abandoned and replaced by another true area of discourse. The difference with fictionalism appears very clearly. Whereas eliminativists hold that the entire relevant discourse must be abandoned in order to be replaced by something else, fictionalists theorize the opposite process, namely, that we must keep the relevant discourse, but, at the same time, we must revise our attitude towards it, which cannot be of believe anymore, but rather must be of simple acceptance or make-believe or any other mental attitude can accommodate a fictionalist perspective. With regard to the fictionalist theories developed by Field and Van Fraassen we can explain this difference in the following way. On the one hand, in his early works, Field does not propose to replace mathematics or to reduce it to a set of sentences only about the concrete world, but rather mathematics is kept unchanged with its usual interpretation that involves the reference to abstract entities. On the other hand, similarly Van Fraassen doesn’t want to eliminate that part of scientific discourse about unobservables and to reduce it to sentences only about the observables, but rather he wants to keep the interpretation of the whole scientific discourse at face value. Hence, according to fictionalists, what changes, in both cases, is not what can or cannot and what

²⁰ Kalderon[2005], p.2.

must or must not be said, but rather what changes is our attitude towards certain areas of discourse, that is not of believe anymore but of make-believe.

Another important distinction is the one between fictionalism and non-factualism. To explain what *non-factualism* is, we can use Kalderon's words:

According to nonfactualism, the sentences of the target region of discourse do not have a truth-evaluable content – they are not genuine representations of a putative domain of fact. [Kalderon 2005, p.4]

In other words, non-factualism can be summarized as the view that the sentences of a specific area of discourse are not used to describe facts. The idea is that there is no reality beyond our sentences formulated within the problematic region of discourse and, as a consequence, there is nothing which the truth and falsity of those sentences depend on. Very often non-factualism can be confused with fictionalism because there is no sharp boundary between the two views. As a matter of fact, non-factualism can be also defined as the view that the correctness of sentences of the relevant area of discourse doesn't amount in being true representations of some facts belonging to reality. However, non-factualists maintain not only that those sentences are not true representations of reality, but also that their truth cannot be assessed, because there is nothing on which to base our evaluation. In fact, we can refer to nothing in the real world in order to evaluate the sentences. The difference between non-factualism and fictionalism is located just at this level. Fictionalists, in fact, don't maintain that sentences of the relevant region of discourse have a non-representational content, but, on the contrary, they maintain that those sentences are genuine representations of certain specific facts of reality. What fictionalists do is claiming that these representations are not true and, as a consequence, that they cannot be believed, but only make-believed. As a matter of fact, according to Field, mathematical sentences represent very well mathematical objects, and the same goes for Van Fraassen who maintains scientific sentences to be genuine representations of the unobservables. For this reason, contrariwise to non-factualists, according to both Field and Van Fraassen, mathematical and scientific sentences are perfectly truth evaluable. In fact, they maintain that their theories have a truth-value – they are or might be false – even if we don't know what this value is.

In conclusion, in order to properly understand what fictionalism is and how it works, it's important to know also what is not and, in particular it's important to distinguish it from both eliminativism and non-factualism in order to prevent fictionalism from being confused with them.

3.1.6 Truth *In* and *About* the Fiction

Another very important issue for a fictionalist approach is the one about the truth of sentences formulated inside and outside the fiction. To this extent it's important to keep in mind the difference between *extra-fictional* and *intra-fictional truth*. This distinction is about what one says *about* the fiction, where, thus, the fiction is the object of the discourse, and what one says *inside* the fiction, where the fiction is the context in which the sentences are put in.

The *extra-fictional truth* is the truth that has to be sought outside the fiction, because it is about how things are in the world. For instance, when we say something like the following:

- (1) Sherlock Holmes is a character created by Sir. Arthur Conan Doyle.

we are talking outside the fiction, that is, we are saying how things are outside the fiction and, in particular, we are saying things about the fiction. As a matter of fact, it is true outside the fiction that Sherlock Holmes is a character created by Sir. Arthur Conan Doyle. If we evaluated (1) within the fiction, (1) would be a false sentence, because inside the story of Sherlock Holmes it is false that Sherlock Holmes is a character created by Doyle, but, instead, it is true inside the story of Sherlock Holmes that, for instance, Sherlock Holmes is a detective.

The *intra-fictional truth*, on the other hand, has to be sought inside the relevant fiction and it does not refer to how things are in the real world, but, instead, it refers to how things are according to the relevant fiction. Let's consider, for instance, the following sentence:

- (2) Sherlock Holmes lives in London in Baker Street 221b.

If we interpret this sentence in an extra-fictional way, it turns out to be false, because in order to evaluate it extra-fictionally we must see to how things are in the world, and, in the reality there is no Sherlock Holmes and in Baker Street 221b there is a bank, not the home of Sherlock Holmes. If, on the other hand, we interpret (2) in an intra-fictional way, that is, we use the fiction of Sherlock Holmes as the context of evaluation for that sentence, then (2) turns out to be true because it is true that inside Conan Doyle's stories of Sherlock Holmes, Sherlock Holmes lives in London in Baker Street 221b.

We can make similar remarks about mathematical fictionalism. As a matter of fact, also with respect to mathematical fictionalism we should distinguish between intra-fictional and extra-fictional truth. Sentences such as the following:

- (3) 2 is the only even prime number.

refer to the fiction of mathematics, that is, they aim at telling us how things are within the story of mathematics, and, hence, they must be evaluated by taken the fiction of mathematics as the context of evaluation. Therefore, (3) is similar to (2) and this means that it is intra-fictionally true. On the other hand, sentences like the following:

- (4) There aren't numbers.

tell us how things are outside the fiction, into the real world. As a matter of fact, (4) tells us that in the world there are no such things as numbers. For this reason, in this case we must look at the extra-fictional truth, similarly to (1). If we evaluate it inside the fiction, (4) will be false, because it is false that inside

the fiction of mathematics there are no numbers. Therefore, (4) cannot be taken as a claim made inside the fiction, but rather a claim about the fiction.

The difference between the two kinds of truth can be made more explicit in the following way: on the one hand, when we talk of extra-fictional truth, we are referring to what we say about the fiction and this corresponds to what philosophers commonly say and discuss when they are inside a seminar room; on the other hand, when we talk of intra-fictional truth, we are referring to what we say inside the fiction and the sentences of this kind are those that are commonly made outside the seminar rooms, that is, they correspond to that kind of discourse that is usually made among a non-specialist audience.

3.2 Varieties of Fictionalism

All those accounts gathered under the label of “fictionalism” have one thing in common: the first rough description of what fictionalism is, namely, that fictionalism is the view that the sentences formulated inside a specific region of discourse are not to be taken literally, because in that case they will be false, but rather they have to be taken as a part of a fiction. Despite this very general character shared by all fictionalist theories, certain distinctions among fictionalist views can be traced within the literature on this topic. These distinctions allow us to further divide fictionalism in different categories. As a matter of fact, as Armour-Garb and Woodbridge explain:

It is important to recognize that what philosophers called “fictionalism” is really a genus of theorizing that we should analyse further into species and varieties. [Armour-Garb and Woodbridge 2015a, p.1]

In what follows we will see the most important differences that we can make inside fictionalism.

3.2.1 Hermeneutic vs Revolutionary Fictionalism

We have seen so far what the main features of a fictionalist account are and what makes it different from the other theories about discourse. Now, we are going to focus our attention on the differences between two specific ways in which we can interpret fictionalism and that correspond to two varieties of fictionalism, namely *hermeneutic* and *revolutionary fictionalism*.

The distinction between hermeneutic and revolutionary fictionalism can be traced back to the distinction between hermeneutic and revolutionary nominalism about mathematics that appears for the first time in 1983 in the works of the American philosopher John P. Burgess.²¹ The first philosopher who applied this vocabulary with respect to fictionalism was Jason Stanley in 2001.²²

Revolutionary nominalism is the nominalistic view that involves a revision of the relevant discourse. The revolutionary mathematical nominalist aims at

²¹ Burgess[1983] and Burgess and Rosen[1997].

²² Stanley[2001].

replacing the current scientific theories – guilty of using terms that lack of reference – with alternative theories that don't involve words referring to mathematical objects. In this way, revolutionary nominalism can be considered a prescriptive approach to the problem because it prescribes what we should do with the problematic discourse. On the other hand, hermeneutic nominalism is a descriptive approach because it does not involve an upheaval and abandon of our current scientific theories, but it merely describes the way in which our scientific theories are actually used. Following Burgess's vocabulary, hermeneutic nominalism is the view that scientific theories, if correctly interpreted, "*already does* dispense with mathematical objects."²³

On the hills of Burgess's analysis, Stanley establishes a difference between hermeneutic and revolutionary fictionalism:

Revolutionary fictionalism would involve admitting that while the problematic discourse does in fact involve literal reference to nonexistent entities, we *ought* to use the discourse in such a way that the reference is simply *within the pretense*. The *hermeneutic fictionalist*, in contrast, reads fictionalism into our actual use of the problematic discourse. According to her, normal use of the problematic discourse involves pretense. According to the pretense, and only according to the pretense, there exist the objects to which the discourse would commit its users, were no pretense involved. [Stanley 2001, p.36 (emphasis in original)]

Therefore, on the one hand, revolutionary fictionalism tells us how we should understand the sentences of the relevant discourse, that is, it tells us that when we use those sentences we ought only pretend to make assertions; on the other hand, hermeneutic fictionalists aim at providing the right interpretation of the relevant fragment of discourse, that is, it tells us that when we use that fragment of discourse we actually don't aim at literal truth, but we only pretend to do so. Hence, in the same way as for nominalism, even for fictionalism the revolutionary version proposes a prescriptive approach to the relevant discourse, whereas the hermeneutic one merely proposes a descriptive approach. As Kalderon well explains the difference:

Let *hermeneutic fictionalism* be a description of a domain of inquiry, and let *revolutionary fictionalism* be a prescription for reforming a domain of inquiry. [Kalderon 2005, p.5 (emphasis in original)]

Hence, on the one hand, hermeneutic fictionalism merely describes the object of its analyses or, in other words, merely tells us how things actually are. As a matter of fact, hermeneutic fictionalism explains that when the speakers use sentences from the relevant region of discourse they are only pretending to make assertions. On the other hand, revolutionary fictionalism maintains that when the speakers utter a sentence from the relevant area of discourse, they wrongly aim at saying a literal truth and, for this reason, revolutionary fictionalism prescribes a revision – or, as the name suggests, a revolution – of the way of talking.

This difference between hermeneutic and revolutionary fictionalism allows us to emphasise another important issue. We can now see what the

²³ Burgess[1983], p. 96 (emphasis in original).

differences between the two fictionalist approaches independently developed by Van Fraassen and Field are. On the one hand, the former is a paradigmatic case of hermeneutic fictionalism because Van Fraassen never proposed a revision of the current scientific theories, but he simply described the behaviour of a speaker towards them. We don't need to believe and assert a theory in order to accept it; it is sufficient that the theory itself has the important feature of the empirical adequacy. On the other hand, Field's theory is a paradigmatic example of revolutionary fictionalism. Field maintains that when we use mathematical discourse we are actually asserting its content and, thus, we actually believe what we say. However, believing and asserting mathematical discourse is a mistake, because it refers to abstract entities and, hence, asserting it, means committing ourselves to the existence of those entities, but, as we know, abstract entities don't exist. Therefore, to repair this mistake, Field proposes the total revision of our attitude towards mathematics and mathematical discourse, in such a way that it is no longer of believe, but rather of simple acceptance.

To sum up the section, the difference between revolutionary and hermeneutic fictionalism can be stated in the following way: on the one hand, revolutionary fictionalist maintains that the fragment of the relevant discourse is, in fact, problematic and we should revise it, that is, we should make it unproblematic through a changing in the aim of the fragment; on the other hand, hermeneutic fictionalist claims the putative problematic region of discourse to be unproblematic, because it actually never had the apparent problematic aim.

What appears very clear is that the two approaches are mutually incompatible and, for this reason, the advantages (and disadvantages) of the former do not correspond to those of the latter, and vice versa. This means that every attack and critique towards the fictionalist perspective must specify the kind of fictionalism which is directed to, because what undermines one specific fictionalist approach does not automatically undermines also the other. However, most of fictionalist philosophers prefer the hermeneutic approach, because, contrary to revolutionary fictionalism, does not totally change our way of talking and does not maintain we are always wrong when we use certain specific areas of discourse. As Armour-Garb and Woodbridge stress again and again:

We favor philosophical fictionalist accounts that attempt to describe the actual functioning of the target discourse rather than propose revisionist strategies for new readings of them. [Armour-Garb and Woodbridge 2015a, p.31]

For this reason, unless otherwise specified, when I'll talk of fictionalism from now on, I'm always referring to the hermeneutic version of the perspective.

3.2.2 Comparative vs Philosophical Fictionalism

Armour-Garb and Woodbridge emphasise that it is possible to distinguish the different fictionalist views also on the basis of the use we make of the

notion of “fiction” into our theorizing.²⁴ This distinction is at the root of the whole work developed by the two American philosophers. In particular, according to Armour-Garb and Woodbridge, two are the possibilities that correspond to two sharply different fictionalist approaches: the comparative and the philosophical fictionalism.

On the one hand, as the name itself suggests, *comparative fictionalism* uses the notion of “fiction” in a comparative way, that is, in order to compare it with the way the objects of the relevant specific theory and correspondent discourse are handled. The comparative fictionalist emphasizes the similarities between certain assertions made within the relevant region of discourse and the assertions made within a work of fiction. As Armour-Garb and Woodbridge puts it:

The similarities are held to underwrite an analogy between (discourse about) the target of theorizing and fiction. [Armour-Garb and Woodbridge 2015a, p.3]

In this case the notion of “fiction” is external to the fictional analysis because we don’t explicitly appeal to it inside the theory. In some cases it is also possible not to appeal to the notion of “fiction” at all.²⁵

Most of the fictionalist accounts we can find within the literature are examples of comparative fictionalism. In particular, some paradigmatic examples of this kind of fictionalism are the early analysis made by Field and the one provided by Yablo.²⁶

On the other hand, contrariwise to comparative fictionalism, in the case of *philosophical fictionalism* the notion of “fiction” has a direct role inside the theory. This is why we say that in philosophical fictionalism the appeal to the notion of “fiction” is internal to the fictionalist analysis itself. The direct use of the notion of “fiction” requires a comparison between the relevant region of discourse and the discourse of the works of fiction, but the main difference with comparative fictionalism is that philosophical fictionalism takes fictionalism a step forward and directly incorporates the fiction within the theory.²⁷ Therefore, while in the case of comparative fictionalism we could, in principle, take the utterances at face value – because the fiction, in this specific case, is external to the utterances –, in the case of philosophical fictionalism this is impossible because the fiction is internal to the utterances, that is, the fiction is incorporated within the utterances themselves and, for this reason, the fiction and the utterances cannot be divided. As Armour-Garb and Woodbridge well explain:

²⁴ Armour-Garb and Woodbridge[2015a].

²⁵ This is the case, for instance, of the theories developed by Field and Van Fraassen.

²⁶ Field[1980] and Yablo[2001].

²⁷ As Armour-Garb and Woodbridge stressed again and again, it is important to keep in mind what the difference between philosophical fictionalism and philosophy of fiction is. Philosophical fictionalism makes use of some elements belonging to fiction in order to solve certain philosophical problems that have nothing to do with fiction. On the contrary, philosophy of fiction investigates issues about fiction itself (such as, for instance, the issues about the existence or non-existence of fictional objects). See Armour-Garb and Woodbridge[2015], pp.4-5.

An account pursuing philosophical fictionalism will start with the thesis that the sentences of D would be semantically infelicitous if they were given a face-value reading. But instead of then making a comparison to fictive or metafictive discourse (however apt a comparison might be), a philosophical fictionalist account instead applies the notion of fiction in explaining how the sentences of D manage to be (or could be made) useful, possibly even to the point of being true. So, on a philosophical fictionalist account, there is no comparison to and modelling on fiction required. [Armour-Garb and Woodbridge 2015, p.18]

The cases of philosophical fictionalism are very rare. In addition to the theory developed by Armour-Garb and Woodbridge we can take as an example of philosophical fictionalism the later analysis made by Field.²⁸

3.2.3 Prefix-Fictionalism vs Pretense-Fictionalism

Another very important aspect on which to base the next distinction among fictionalist theories is the one about the application of the notion of “fiction” within the fictionalist account. In particular, I’m wondering how that application is made. To this regard, it is possible to distinguish the prefix-fictionalism from the pretense-fictionalism, which depends on the theory of the fiction we decide to base our fictionalist account on. In fact, if we decide to use the philosophy of fiction of David Lewis,²⁹ then our account will be a case of *prefix-fictionalism*; if, on the contrary, we decide to rely on Kendall Walton’s philosophy of fiction,³⁰ then our account will be a case of *pretense-fictionalism*.

According to Lewis’ theory of fiction, all the metafictive sentences about a fiction are semantically infelicitous if they are given a face value reading, because certain terms involved in them are lacking of reference. This is the case, for instance, of the following sentence:

- (5) Sherlock Holmes is a detective.

Since the name “Sherlock Holmes” refers to a fictional entity and since fictional entities don’t exist, then “Sherlock Holmes” is denotationless in Lewis’ view. Therefore, according to Lewis’ account, metafictive sentences, such as (5), are not true.

The central idea in the Lewisian analysis is that in sentences like (5) there is a tacit prefix specifying the fiction of reference, which is what allows establishing the context of evaluation for the sentence. The sentence (5) involves a tacit prefix of the kind “according to the fiction of Sherlock Holmes” and, more in general, all metafictive sentences involve a tacit prefix of this kind: “according to the fiction F”. In this way, metafictive sentences must be evaluated on the adherence of what they say to the fiction they belong to. Hence, in Lewis’ view, (5) is an abbreviation of the full sentence that involve the right tacit prefix which specify the relevant fiction and that can be stated as follows:

²⁸ Field[1989].

²⁹ Lewis[1978].

³⁰ Walton[1990 and 1993].

- (6) According to the fiction of Sherlock Holmes, Sherlock Holmes is a detective.

It's now pretty clear how the varieties of prefix-fictionalism work on the basis of Lewis' theory of fiction. A paradigmatic example of prefix-fictionalism is Field's theory of mathematical discourse. According to Field, mathematical objects don't exist and, for this reason, mathematical sentences are semantically infelicitous if they are given a face value reading. Thus, the right way to interpret the sentences of pure mathematics, such as the following:

(7) $2+2=4$

and of applied mathematics, such as the following:

- (8) The Martian moons are 2.

is not literally, but rather fictionalistically. In particular, it is necessary to postulate the existence of a tacit operator functioning as a prefix such that it makes both (7) and (8) true on the basis of the adherence of what they say to the fiction established by the prefix. The tacit prefix for the two sentences is something like "according to the fiction (story) of mathematics" and it is without doubts true that $2+2=4$ according to the story of mathematics (and the same goes also for (8)).

It's important to explain now what the relation between the unprefix sentences and the correspondent sentences with the relevant prefix is. They cannot be conceived as synonymous, because if this were the case, then the fictionalist prefix would be redundant and there would be no difference in the truth-values of the two sentences. The idea, instead, is that we have to distinguish between two different meanings associated with the sentence: the *semantic meaning*, namely, the literal meaning of the sentence; and the *speaker meaning*, namely, the real meaning of the sentence, that is obtained by affixing the relevant fictionalist prefix to the sentence.³¹

Pretense-fictionalism is based upon the theory of fiction developed by Walton. Walton maintains that the fiction acting in sentences like (5) is very similar to that involved in children's games of make-believe. The games of make-believe are characterized by the use of the so-called *props*, that are objects belonging to the real world that makes the propositions made within the world of the make-believe true. In other words, *props* are supports that allow making fictionally true certain propositions that are literally false. For instance, when in a game of make-believe played by children the round pieces of mud are pies, then we can qualify the round pieces of mud as *props* for that specific game of make-believe. Now, there are two possibilities: either the *props* operate as supports for the game of make-believe and the focus of our attention is the content of the game of make-believe, so, in this case, we say that the game of make-believe is *content-oriented*; or the game of make-believe is used as an instrument for the comprehension of the *props* themselves – and, hence, the

³¹ We talked about this difference also in section 3.1.1.

props are the real aim of our engaging in the game of make-believe – and, in this case, we say that the game of make-believe is *prop-oriented*. In the latter case, knowing what is fictionally true allows us to know what is true in the reality, that is, looking at how things are inside the fiction allows us to know how things are outside it, in the reality. Let's consider, for instance, the following sentence:

(9) Crotone is in the arch of the Italian boot.³²

This sentence invites us to see Italy as a boot. Let's suppose we want to know where Crotone exactly is. The game of make-believe we are engaged in is what allows us to know that “Crotone is in the place it would need to be, to make it acceptable in the game that Crotone is in the arch of the Italian boot.”³³ Knowing where Crotone is in the fiction with respect to the boot allows us to know where Crotone is in the reality with regard to the map of Italy. Hence, this is a paradigmatic example of prop-oriented game of make-believe. More in general metaphors are all examples of prop-oriented game of make-believe. On the other hand, the content-oriented game of make-believe is typical of the fictional world of novels, movies, paintings, etc.

The other important elements characterizing the game of make-believe are the base fictions that are stipulated at the beginning and the principles of generation that determine which new fictions, besides the base ones, are at work in the game of make-believe. The principles of generation use the *props* to establish the new fictions, and, for this reason, we can say that the principles of generation connect how things are inside and outside the game of make-believe with respect to objects belonging to reality, i.e. the *props*. In other words, the principles of generation are the rules establishing what is legitimate and what is not within the game of make-believe. By recalling the previous example, the principles of generation establish that it is true according to the pie story that x is a cook if and only if x makes round pieces of mud.

Therefore, when we use sentences like (5) – i.e. “Sherlock Holmes is a detective” – we are implicitly subscribing the rules of the game of make-believe of the stories of Sherlock Holmes. Being engaged in the game of make-believe means subscribing its rules, which, in turn, means accepting that (5) is true and its negation is false.

To sum up, the difference between Walton and Lewis' accounts is the following: on the one hand, the latter involves an operator specifying the relevant fiction and that plays the role of a tacit prefix that, by so doing, allows us to say that (5) is an abbreviation of the full sentence (6); on the other hand, Walton takes the unprefixated sentences as primary, since (5) is taken by Walton as a *prop* that, through the base fictions and the principles of generation, allows us to indirectly assert what we would directly assert with a sentence of the form “the Sherlock Holmes fiction (story) portrays someone named ‘Sherlock Holmes’ who is a detective.”

If we consider the sentences of pure and applied mathematics, (7) and (8), what we obtain with this account is that if they are given a face value reading, they are plainly false, but when we use the number-terms in them as *props*, that

³² Walton[1993].

³³ Yablo[2014], p.172.

is, as supports in the game of make-believe of mathematics that is ruled by the stipulated fictions and the specific principles of generation, the two sentences turn out to indirectly convey what the following sentences directly assert. In particular, (7) indirectly conveys what follows:

- (9) The mathematical fiction (story) is such that there are numbers and that $2+2=4$.

And, (8) indirectly conveys what follows:

- (10) The mathematical fiction (story) is such that there are numbers and that the Martian moons are 2.

Hence, when we use fictional discourse we are using a *prop*-oriented make-believe, because we are not interested in how things are within the fiction, but rather we are interested in the reality, and we use fictional discourse in order to know how things are in the real world. When we say “the Martian moons are 2” or, to use the famous example of Walton, “Crotona is at the arch of the Italian boot”, we are not interested respectively in numbers and in the boot, but rather the fiction is an instrument useful to know how things are in the reality. In the former case, we are using the fiction to know how things are with regard to the moons of Mars and, in the latter case, we use the fiction to know how things are about the geography of Italy and in particular about the location of Crotona on the map of Italy.³⁴

A paradigmatic example of fictionalism based on Walton’s pretence theory is the one developed by Armour-Garb and Woodbridge about several areas of discourse (propositions-talk, existence-talk and truth-talk, which our analysis is particularly interested in),³⁵ and the one developed by Yablo about the discourse on possible worlds.³⁶

It is worth noting that being a prefix-fictionalist or a pretence-fictionalist is an independent issue from being a hermeneutic or a revolutionary fictionalist. Let’s assume we wish to defend a prefix-fictionalist’s view. If we prefer a hermeneutic perspective, what we are claiming is that when we say something like “The Martian moons are 2” what we really mean is something of the form “According to the fiction of mathematics, the Martian moons are 2”. On the contrary, his revolutionary counterpart maintains that with an utterance of the former we are making a wrongly statement and that we must revise our way of talking and use the latter sentence from now on. In the light of these considerations, we can conclude that the different varieties of fictionalism can be combined with each other and this is exactly what actually happens.

3.2.4 Yablo’s Distinction

According to Yablo, it is possible to distinguish four different varieties of fictionalism “according to the various accounts one might give of ‘advancing in

³⁴ Yablo[2001].

³⁵ Armour-Garb and Woodbridge[2015a].

³⁶ Yablo[1996].

a fictional spirit’.”³⁷ More in details, the four different varieties of fictionalism pinpointed by Yablo are the following: instrumentalism, meta-fictionalism, object-fictionalism and figuralism. Let’s see now what are the main differences between them.

According to *instrumentalism*, when the speaker utters a sentence from the fictional discourse, she is not really asserting anything but rather she is only pretending to assert something. Why do we have to pretend to assert something? Which is the purpose of this “make as if”? The purpose is to say something that has nothing to do with the fictional discourse. In other words, the reason why we choose to pretend to assert something is that fictional discourse, albeit it is false, is somehow useful to us. As Yablo puts it:

Our reason for making as if we are doing these things [i.e. asserting and/or believing that S] is that it serves some larger purpose. Making as if S enables us simplifies our theory, or shorten proofs. Someone who stops here – someone with no story to tell about what we are “really doing when making as if S, and why that would be a sensible thing to do – I will call an *instrumentalist fictionalist*, or simply *instrumentalist*. [Yablo 2001, p.74 (emphasis in original)]

The main features of instrumentalism, hence, are essentially two: on the one hand, the account requires the speaker to solely pretend to assert; on the other hand, an alternative account of what the speaker does is not presented. This is the less developed and, for this reason, the less interesting form of fictionalism.

Moreover, instrumentalism faces a bunch of problems, among which the most significant are *the problem of the real content* and *the problem of correctness*. The former is about assertion and the pretense of assertion. Instrumentalists are simply not interested in the content of a sentence. According to instrumentalism, when we utter a sentence like (8), i.e. “The Martian moons are 2”, we are not really asserting that the number of the Martian moons is 2, but we only pretend to make such an assertion. The problem is that there seems to be something I’m really and genuinely asserting with an utterance of (8) and instrumentalism doesn’t explain what it is. In other words, instrumentalist doesn’t explain what the real content of sentences like (8) is. The latter problem is the problem of correctness, that is, the problem that by subscribing the instrumentalist perspective we cannot account for the difference between an utterance of (7) – i.e. “ $2+2=4$ ” – and an utterance of “ $2+2=5$ ”. By intuition we would like to say that an utterance of “ $2+2=4$ ” is correct and an utterance of “ $2+2=5$ ” is incorrect. The problem is connected with the previous one, because an utterance is correct if and only if its real content is true, but, as we saw, instrumentalists neither provide the account with a way for establishing what the real content of an utterance is nor, for this reason, they provide a way for distinguishing an utterance of “ $2+2=4$ ” from an utterance of “ $2+2=5$ ”, that is, they do not provide the account with a way for characterizing the former as correct and the latter as incorrect.

The second variety of fictionalism pinpointed by Yablo is the so-called *meta-fictionalism*, namely, the view that the speaker is really asserting something, and, in particular, she is asserting that things are so and so according to the

³⁷ Yablo[2001], p.74.

relevant fiction. In other words, according to meta-fictionalism, when we talk inside the fiction, we are really talking about the content of the fiction. As Yablo puts it:

It [i.e. meta-fictionalism] says that in making as if to assert that S, one is really asserting that S is the right kind of thing to make as if to assert: the quasi-assertion game one is involved in endorses the quasi-assertion that S. [Yablo 2001, p.76]

This is the kind of fictionalism endorsed by Field in 1989 with his work *Realism, Mathematics and Modality*. In this book Field advances the idea that when we say “the Martian moons are 2”, what we are really saying is “according to the fiction of mathematics, the number of the moons of Mars is two”.

Also meta-fictionalism faces a bunch of problems, the most important of which is of phenomenological character. When we say “The number of starving people is very large”, according to meta-fictionalism, what we are really saying is something like this “according to the fiction of mathematics, the number of starving people is very large”, but it is clear that it is not correct because the content of our sentence is not mathematics, but rather people, and meta-fictionalism fails in accounting for that. A second problem affecting meta-fictionalism is the modal problem. According to meta-fictionalism, mathematical statements could have been different. More specifically, from the one hand, ordinarily speaking, mathematical statements such as “ $2+2=4$ ” are necessary true – that is, $2+2$ is necessary equal to 4 – but, on the other hand, “it could (perhaps) have been otherwise that $2+2=4$ according to standard math” that, in other words, means that “it is not *a priori* that standard math turned out the way it did”.³⁸ Hence, there seems to be a conflict between the ordinary use and the fictional use of mathematical sentences: according to the ordinary use, when we assert a mathematical sentence we are asserting something that is *a priori* necessary but, it is neither *a priori* nor necessary that, according to the story of mathematics, things are so and so.

Object-fictionalism is the third variety of fictionalism we are going to analyse. According to object-fictionalism, the speaker is really asserting something and what she is asserting is that the world is in certain conditions, that is, those conditions it needs to be in in order to make it true in the relevant fiction that things are so and so.³⁹ Hence, object-fictionalism can be seen as the view directly opposite to meta-fictionalism, because according to object-fictionalists what we are asserting with a fictional sentence, i.e. the real content of a fictional sentence, is something non-fictional about the real world. Therefore, by using the same example as before, when we say “the Martian moons are 2” we are actually saying how things are in the reality and not in the fiction of mathematics. The fiction is only a useful tool in the achievement of this goal. We can restate the issue by using the most effective words of Yablo:

When the words “the number of apostles is twelve” come out of my mouth, I am relying on the number-fiction. It is thanks in part to that fiction that my utterance is a way of saying that there are twelve apostles.

³⁸ Yablo[2001], p.76 (emphasis in original).

³⁹ Eklund[2015].

Chapter 3

Again, though, relying on a fiction is one thing; talking about it is another. The fiction functions as medium and not message. [Yablo 2001, p.76]

Hence, in this case, the fiction and the non-existent objects belonging to the fiction work as instruments, or, as Yablo calls them, as *representational aids*.⁴⁰ Eklund says what follows about fictionalism about mathematical discourse:

The characteristic objects of the discourse, the numbers, are mere aids we use to make utterances about how things stand in the real world. [Eklund 2015]

The advantage of this perspective over the others is twofold. In the first place, it allows us to solve the phenomenological problem faced by meta-fictionalist perspectives. When we say “the number of starving people is very large”, we are not talking about mathematics, as it turns out if we endorse a meta-fictionalist view, but rather we are talking about people, i.e. about reality. In fact, from an object-fictionalist point of view, what we are saying with an utterance of that sentence, i.e. its real content, is something like “there are many starving people” and, thus, we are keeping the real subject matter of the sentence that, we should stress, is not the fiction of mathematics but people. Moreover, object-fictionalism helps in the solution of the modal problem. In fact, as we saw, what makes a sentence true within the relevant fiction is its real content and, if we use the previous example, what make the sentence “ $2+2=4$ ” true is the fact that there are two F’s and two G’s, which, as a result, gives us four F-or-G’s.

Nevertheless, also object-fictionalism faces a major problem, which inspires Yablo in the development of his own fictionalist account in order to solve it. Yablo calls the problem *the bomb*.⁴¹ The bomb directly arises from one of the most peculiar features of the fictionalist account, namely, its conditions of quasi-assertion. We know that a sentence is quasi-assertible if and only if it is true within the relevant fiction, and it is true within the relevant fiction if and only if its real content obtains. If we take the fiction of applied mathematics to be the relevant fiction, we obtain the following argument:

- (1) “ $\#(K's)=n$ ” is quasi-assertible iff according to the story of applied math, $\#(K's)=n$.
- (2) “ $\#(K's)=n$ ” is quasi-assertible iff there really are n K’s.
- (3) According to the story of math, $\#(\text{even primes})=1$.⁴²
- (4) “ $\#(\text{even primes})=1$ ” is quasi-assertible. [from 1 and 3]
- (5) There really is an even prime number. [from 2 and 4]

However, of course no mathematical fictionalist would accept (5), because, on the contrary, she would actually accept the negation of (5), namely, there is no

⁴⁰ Yablo[2001], p.81.

⁴¹ Ibid. pp.78-80.

⁴² 1 and 2 follows from the conditions of quasi-assertion, and 3 is an obvious truth.

even prime number. Hence, from certain premises commonly accepted by every fictionalist follow a conclusion that no fictionalist would accept. As a matter of fact, mathematical fictionalism holds that numbers don't exist but from the fictionalism theses follows the awkward conclusion that numbers do exist, because there is at least one number, i.e. an even prime number.

The problem with this account is that even if we reverse the argument, that is, even if we use as a premise the negation of the conclusion of the previous argument, which is also one of the fundamental theses of mathematical fictionalism, things don't go better. Let's see how the argument proceeds:

- (14) There really are no numbers.
- (14) “ $\#(\text{numbers})=0$ ” is quasi-assertible.
- (14) According to the story of math, $\#(\text{numbers})=0$. [from 1 and 2]

However, also this conclusion goes against the fictionalist thesis. In fact, it is false that according to the story of math, the number of numbers is 0. What is true, according to fictionalism, is that the number of numbers is 0 if taken at face value, but it is obviously false that this is the case according to the story of math.

Furthermore, this latter result is even worse than the one of the previous argument. The reason is well explained by Yablo:

The results we are getting seem in fact to be worse than false. Take “the number of numbers is 0”. That has no chance of being true, because it is self-refuting. [Yablo 2001, p.79]

It is, in fact, intuitive that if the number of numbers is 0, then there is at least one number, namely 0. This, however, means that the number of numbers is not 0, but it is at least 1. This is why we say that the sentence “the number of numbers is 0” is self-refuting.

The last variety of fictionalism we are going to analyse is the one defended by Yablo himself and is called *figuralism*.⁴³ According to Yablo, a given sentence from the relevant discourse is associated with two different contents: its semantic content, that is the literal one, and the fictionalist does not believe it, but only make-believe it; and the figurative content, that is the real content of the sentence, which the fictionalist really believes. Figuralism maintains the main idea of object-fictionalism, namely, that numbers are representational aids but, contrariwise to object-fictionalism, figuralism does not stop at this point and takes this idea one step forward by taking into account also all those situations where numbers function as things represented. To use an example from Yablo, the difference between the two uses is the same that there is between the two occurrences of the name “butterflies” in the following sentences:

- (11) I had butterflies in my stomach.

⁴³ Yablo[2001] calls this kind of fictional theory also with the name of *reflexive fictionalism* in order to emphasize the reflexive use of certain notions involved into the account.

and

(12) The butterflies were splatted all over the windscreen.

In (11), the name “butterflies” is used in a figurative way to suggest that I was worry about something and, therefore, it works as a representational aid. In this case, (11) must not be given a face value reading, but rather fictionalistically. By contrast, in (12), we are really talking about butterflies, the bugs, and, therefore, the name “butterflies” here works as thing represented and (12) must be read at face value.

If we apply this account to mathematics, we obtain what follows. When we make an utterance of a sentence like (8) – i.e. “the Martian moons are 2” – we are using numbers as representational aids in order to talk of the concrete world. But let’s consider the following sentence:

(13) There are numbers.

With an utterance of (13) we are not using numbers as an instrument in order to talk about reality, but rather we are actually talking about numbers, that is, we are really making an assertion about numbers. In this case numbers function as things represented.

However, there is another possibility to take into account. This is the case of mixed sentences such as the following:

(14) The number of even prime numbers is 1.

In (14) numbers function both as representational aids (in the first occurrence of the name, i.e. “the *number*...”) and as things represented (in the second occurrence of the name, i.e. “... even prime *numbers*...”). This allows us to provide an answer to the problem of the bomb. The problematic sentence of the bomb, i.e. “the number of numbers is 0” turns out not to be self-defeating on this account, and, on the contrary, it makes perfectly sense if we interpret it in the way figuralism suggests. The first occurrence of “number” must be interpreted in a figurative sense, namely as a representational aid, while the second occurrence must be interpreted at face value, namely as thing represented. In this way, there is no conflict between the two occurrences and the sentence is not self-defeating, and in this way the explosion of the bomb is avoided and it is not a problem for fictionalism anymore.

3.3 Motivations for Fictionalism

There is a whole series of arguments in favour of a fictionalist perspective, we already met and faced some of them in the chapter. In this context, I shall just address and elaborate those arguments we have simply glazed over and I shall mention and analyse some of the most famous among the others. The aim is to give an idea of what the advantages in the endorsement of this kind of perspective can be.

3.3.1 Ontological Economy

The main reason why someone might want to adopt a fictionalist perspective of a certain area of discourse is, as I already stressed, tied to motivations of ontological economy. In particular, what encourages most philosophers in the endorsement of a fictional view of a specific region of discourse is the need to avoid an inconvenient ontological commitment towards entities that are the putative references of sentences belonging to the relevant region of discourse. The reasons to avoid this ontological commitment are manifold, but most of them hinge on considerations of ontological economy. One may wonder, for instance, what follows: how do we have to consider mathematical discourse provided that the words involved in its sentences actually refer to nothing? How do we have to understand sentences like “ $2+2=4$ ” and “the Martian moons are 2” provided that the objects they refer to are abstract objects and abstract objects don’t exist?

Fictionalists provide an answer to all these questions. The value of this answer is that it doesn’t require the appealing to the problematic alternative of paraphrases, which maintains that the sentences of the relevant discourse that appear to express certain propositions, actually express certain other propositions. It seems that mathematical sentences are talking about abstract objects, but they actually are not. A sentence like “ $2+2=4$ ” if taken at face value conveys an ontological commitment to numbers and this is why the advocates of the strategy of paraphrases propose to consider the following true paraphrases of that sentence: “if there were such things as numbers, then $2+2=4$.” On the one hand, fictionalists agree with the advocates of the strategy of paraphrases in what they say about the existence of numbers, namely, that they don’t exist; but, on the other hand, they disagree with the idea that the sentence “ $2+2=4$ ” must not be taken literally. Therefore, according to fictionalists, “ $2+2=4$ ” must be taken literally and, at the same time, they maintain that mathematical objects don’t exist. The conclusion is that, according to fictionalists, “ $2+2=4$ ” is literally false.

Let’s now come back to the questions raised a moment ago: how can we consider those sentences? What do we have to do with the problematic discourse involving them? The fictionalist answer is that those sentences must be considered in a similar way to those we utter within a story of fiction. In particular, as we already saw, the sentence “ $2+2=4$ ” turns out to be true according to the fiction of mathematics. Therefore, fictionalism provides a viable alternative to the strategy of paraphrases for all those philosophers that wish to make the ontological economy a priority.

Moreover, fictionalism turns out to be useful in order to avoid an inconvenient ontological commitment also for those philosophers who wish to avoid certain epistemological worries that traditionally are imputed to the Platonist views of mathematical objects. The problem is the following: one could wonder how it is possible to have knowledge of mathematical objects (and the same goes for any other problematic entity such as, for instance, propositions, possible worlds, and so on) provided that we don’t enter in causal relation with them in any way.⁴⁴ If mathematical objects existed, they should

⁴⁴ The objection, of course, presupposes that to have knowledge of an object we need to be in causal relation with such an object.

have a role in the acquisition of mathematical knowledge, but this seems not to be the case. Of course, such a problem disappears for all those philosophers who assume a nominalistic perspective of mathematics, according to which mathematical objects don't exist.

Besides these reasons tied to the problem of the ontological parsimony, there are also other arguments in favour of fictionalism that it is worth noting. Take a look at what they say in details in next sections.

3.3.2 The Paradox of Existence

Another argument connected to the issue of the ontological commitment that is worth noting is due to Yablo and is commonly known with the name of *paradox of existence*.⁴⁵ If, on the one hand, the issue about the existence or non-existence of abstract entities is very complex, on the other hand, there seems to be some trivial argument in favour of the idea that abstract entities really exist. The main argument leads us to admit the a priori existence of abstract objects on the basis of one or more premises that are, in turn, a priori and/or of undisputed empirical facts. If we consider, for instance, the sentence “ $2+2=4$ ”, we have no empirical evidence in favour of this equivalence and, for this reason, we say that its knowledge is a priori. Therefore, when we say something like “ $2+2=4$ ”, it seems intuitively correct to think that what follows is a perfectly valid argument: “if $2+2=4$, then there is a number that added to 2 yields 4. This means that there is something that is a number and, hence, as a consequence, we can say that there are numbers.” The same reasoning applies for other problematic entities such as, for instances, properties. When we utter an apparently innocent sentence like “the tomato is red”, we can build an argument for properties very similar to the previous one about numbers. The argument is the following: “if the tomato is red, then the tomato has the property of redness. This means that there is something that is a property and, hence, as a consequence, we can say that there are properties.”⁴⁶

If we are nominalists about numbers or properties (or about any other problematic entity), how can we solve the paradox of existence? The solution provided by Yablo is to assume a fictionalist approach of the discourse that involves words whose reference is those problematic entities. As a matter of fact, since we are inside a philosophy's seminar room the issue about the existence of those entities is very complex because the relevant sentences must be taken literally. If, on the contrary, we use those sentences outside the philosophy's room, i.e. in an ordinary place such as the street, then the issue becomes simpler because in this case those sentences must be taken as made inside a fiction and they can be evaluated only as fictionally true and not literally. When an ordinary speaker claims something implying the existence of numbers, she actually doesn't want to say anything about numbers, but rather she wants to indirectly express something about the concrete world. For this reason, that claim turns out to be true according to the relevant fiction, that is, the fiction of mathematics. On the contrary, the same claim uttered within a

⁴⁵ Yablo[2000].

⁴⁶ Yablo[2000] shows that it is possible to construct similar arguments for all those problematic entities whose existence is questioned, namely, possible world, events, countermodels, sets, propositions and so on.

philosophical debate about the existence of numbers acquires completely another meaning; its real content, in this case, corresponds to its semantic content and, hence, the claim must be taken at face value. To say that by using the vocabulary of section 3.2.6, if we are inside a philosophy's room, we should evaluate the sentence in an extra-fictional way, whereas if we are speaking to ordinary people, the sentence in this context must be intra-fictionalistically evaluated.

3.3.3 The Oracle

The most famous argument in favour of a fictionalist perspective is known as the *argument of the Oracle*.

The following mental experiment is due to Burgess and Rosen⁴⁷ and it was later developed and enriched by Yablo.⁴⁸ Let's say you have the possibility of appealing to the Oracle of Philosophy and let's say the Oracle is omniscient and you can ask him whatever you want. Following Burgess and Rosen, we can imagine the dialogue will proceed and develop in the following way:

Finally, after years of waiting, it is your turn to put a question to the Oracle of Philosophy. So you humbly approach and ask the question that has been consuming you for as long as you can remember: "Tell me, O Oracle, what there is. What sorts of things exist?"

To this the Oracle responds: "What? You want the whole list? Look, I haven't got all day. But I will tell you this: everything there is is concrete; nothing there is is abstract. Now go away and don't bother me." [Burgess and Rosen 1997, p.3]

Despite the grumpiness and rudeness of the Oracle, you firmly believe in the truth of his claims and, therefore, you firmly believe that only concrete entities and no abstract entity really exist.⁴⁹ Hence, you leave the Oracle with the awareness that nominalism about mathematics (and nominalism in general) is the right perspective to assume in philosophy of mathematics.

Many questions now arise, such as the following: what happens with regard to the discourse about abstract entities? Thanks to the Oracle we know there are no abstract entities and, hence, we know the discourse about them to be denotationless. How do we have to behave towards that region of discourse? Do we have to keep talking in the same way as before or we must abandon the problematic areas of discourse in order to change our way of talking? Thanks to this new awareness, do we keep saying things like "the Martian moons are 2" even if there is nothing which the word "2" refers to, or not? Would we, perchance, say to mathematicians and scientists that they should stop using sentences like "2+2=4" because they involve denotationless words? Would something change for mathematicians and their works, and also for ordinary people, when the knowledge of this truth about abstract entities

⁴⁷ Burgess and Rosen[1997], p.3.

⁴⁸ Yablo[2000].

⁴⁹ The argument of the Oracle works not only for abstract entities, but also for other problematic objects such as moral facts. For this reason the argument of the Oracle works for all kind of fictionalism about any region of discourse and, in this way, it provides a strong reason in favour of fictionalism as a *genus*.

comes into play or, maybe, they will keep talking as always before? And, finally, if we keep talking as always before, does our way to understand mathematical sentences change?

These are some of the issues arising from the mental experiment of the Oracle. The swiftest and most intuitive answer to those questions is that even if the Oracle told us that numbers don't exist – and, more in general, that all abstract entities don't exist –, nothing would change in our way of talking. We would keep using the same regions of discourse and the same sentences as before. Moreover, also our way of understanding and interpreting those sentences wouldn't change. After all, when we use the sentence “the Martian moons are 2”, we are talking about Mars and its moons, and not about numbers, or mathematical objects in general. The fact that the latter don't exist doesn't change our way of talking about concrete objects, i.e. about the world. In other words, the use we made of mathematical discourse doesn't change because we use the sentences from this region of discourse as representational aid, that is, we use it not to talk about mathematical objects, but in order to talk about something else, i.e. concrete objects.

All these comments provide strong reasons in favour of fictionalism and, in particular, in favour of hermeneutic fictionalism. Indeed, to be more accurate, the argument of the Oracle goes in favour exclusively of hermeneutic fictionalism, and not of its revolutionary counterpart. The latter, remember, is a prescriptive variety of fictionalism and, for this reason, presupposes the full reform of the problematic area of discourse, but what the mental experiment of the Oracle shows is that this would not be our behaviour. The fact that our attitude towards the problematic regions of discourse would remain unchanged suggests that there is no need to make a revision of that discourse, in contrast to what revolutionary fictionalist claims. The hermeneutic fictionalist, instead, simply describes our way of talking, that is, simply says that from the beginning when we use the relevant fragment of discourse we are not aiming at literal truth, but rather we are only pretending to do so and, thus, hermeneutic fictionalist merely describes how things are and she does not say how things should be. This really seems to be the attitude we would adopt if the Oracle told us that abstract entities don't exist, that is, we would keep using the fragment of discourse involving abstract entities in the same way as before, with the awareness that we were always engaged in a game of make-believe, or, more generally, in a fiction.

Besides being an argument in favour of hermeneutic fictionalism, the Oracle argument is strictly linked to another argument against revolutionary fictionalism, namely, the *argument of immodesty*.⁵⁰ It's important because it accuses revolutionary fictionalism of immodesty. If we found out that mathematical objects don't exist, for instance thanks to the omniscient Oracle, and that, for this reason, mathematical-talk is false if taken at face value, then revolutionary fictionalism would force us to change our attitude towards mathematical claims, that is, according to revolutionary fictionalists we should stop believing mathematical sentences and, instead, we should start believing them to be true only according to the relevant fiction. This means that we should stop believing “ $2+2=4$ ” and we should only believe the stronger sentence “according to pure

⁵⁰ Burgess[2004], p.30.

mathematics, $2+2=4$.” The problem with this proposal is the following: how can philosophers dare to say to mathematicians they are wrong about their studies and works and they should change their attitude towards the objects of their works? Almost everyone is willing to admit that this would be unacceptable. As David Lewis emphasizes:

To reject mathematics for philosophical reasons would be absurd. [...] How would you like to go and tell the mathematicians that they must change their ways, and abjure countless errors, now that philosophy has discovered that there are no classes? [Lewis 1993, pp.14-15]

Therefore, doing so would make philosophers too much presumptuous and arrogant. They should, instead, have the modesty of not spreading through fields not covered by their mandate.

Hence, even if we found out that mathematics is false, what we should do is assuming a fictionalist perspective and, more specifically, a fictionalist perspective of hermeneutic type, because it merely aims at describing the behaviour the speakers already have towards mathematical sentences and it does not propose a radical change in their attitudes.

3.4 Truth-Theoretic Fictionalism

Fictionalism about truth (also called *alethic fictionalism* or *truth-theoretic fictionalism*) is a branch of fictionalism that has been very recently developed mainly thanks to the works of Armour-Garb and Woodbridge,⁵¹ Beall,⁵² Burgess⁵³ and Woodbridge alone.⁵⁴ How we can easily guess, fictionalism about truth is the view that the discourse about truth, that is, the sentences formulated by using the truth predicated and/or the falsity predicate, if given a face value reading are false, whereas they are true inside the relevant fiction of truth. More in details, fictionalism about truth is the idea that, although there is no substantive property of truth, the discourse involving the truth predicate can still be used. In other words, although the truth-talk is literally false, it is still useful because it allows the expression of things that have nothing to do with truth and that we could not express otherwise. As Armour-Garb and Woodbridge put it:

A slogan for our SPIF [i.e. semantic pretense-involving fictionalism] account of truth-talk is the claim “Truth is pretense,” the idea being that while we talk as if there were a property of truth, this fragment of discourse operates with complete indifference as to whether there really is any such property. [Armour-Garb and Woodbridge 2015, p.130 (emphasis in original)]

⁵¹ Armour-Garb and Woodbridge[2014 and 2015a].

⁵² Beall[2004].

⁵³ Burgess[2007].

⁵⁴ Woodbridge[2005].

This means that we talk as if there is a substantive property of truth, but actually we pretend such a property to exist merely for expressiveness's sake.

In this section I'm going to focus mainly on the fictionalist view developed by Armour-Garb and Woodbridge. The reason is that, unlike other views, the argument presented by Armour-Garb and Woodbridge together does not aim at providing a specific fictionalist account, but it merely introduces a very general position on this matter that allows the philosopher who wants to embrace a fictionalist perspective of truth-talk to choose among a wide range of options.⁵⁵

The argument developed by Armour-Garb and Woodbridge I'm referring to is included in the paper *From Mathematical Fictionalism to Truth-Theoretic Fictionalism*⁵⁶ and then it is used to further develop their fictionalist theory in their last book published very recently, in 2015.⁵⁷ The two authors wish to support a fictionalist perspective about truth-talk and, more in details, they wish to prove that a deflationist philosopher should subscribe such a fictionalist perspective. The aim, thus, is to highlight the reasons why a theory that claims truth to have no underline nature, but, on the contrary, claims that all we can say about truth is limited to the instances of the T-schema, should embrace a fictionalist perspective about truth-talk. In other words, the two philosophers wish to clarify the reasons why a deflationist about truth should take truth as an element of fiction and, to be more precise, the discourse about truth as a discourse about a specific kind of fiction, i.e. the fiction of truth.

The argument used by Armour-Garb and Woodbridge proceeds by analogy with the argument developed by Stephen Yablo in *The Myth of the Seven*⁵⁸ in order to justify his fictionalist view about mathematical discourse.⁵⁹ The crucial point of the argument of Armour-Garb and Woodbridge is the following: if Yablo's argument, according to which nominalistic philosophers of mathematics should assume a fictionalist perspective of mathematical discourse, works, then the argument according to which deflationists about truth should assume a fictionalist perspective of truth-talk will work as well. The reasons are the same for both accounts, that is, the motivations for a deflationist in assuming a fictionalist perspective of truth-talk are the same as for a nominalist in assuming a fictionalist perspective of numbers-talk.

Let's now see very shortly Yablo's argument for numbers-talk and then have a look at the analogies among it and the one about truth-talk developed by Armour-Garb and Woodbridge.

3.4.1 Fictionalism about Mathematical Discourse

To make things easier we might say that mathematical fictionalism is the view that the following premises hold:

⁵⁵ I'm going to analyse some among the other fictionalist theories in the final chapter.

⁵⁶ Armour-Garb and Woodbridge[2014].

⁵⁷ Armour-Garb and Woodbridge[2015a].

⁵⁸ Yablo[2005].

⁵⁹ Actually, Yablo wishes to justify his own specific fictionalist perspective of number-talk, i.e. figuralism, but the first part of his argument can be used to justify mathematical fictionalism in general. This is the part Armour-Garb and Woodbridge refer to in their works.

Truth-Theoretic Fictionalism

- a) Mathematical objects, in so far as they are abstract objects, don't exist (nominalism);
- b) Mathematical sentences are, in fact, about mathematical objects.

and from these premises the following conclusion follows:

- c) Mathematical sentences, along with mathematical theories, if taken at face value are false.⁶⁰

It's clear, thus, that at the root of a fictionalist theory about mathematical discourse there must be the assumption of validity of the two premises and, in particular, of the first one that is nothing but the nominalistic thesis about mathematical objects.

According to Yablo, mathematical fictionalism is the best perspective a nominalist could ever assume because it allows the explanation of why we can use mathematical discourse despite the objects which that discourse refers to don't exist and, as a consequence, why the sentences belonging to it are literally false. In other words, fictionalism about mathematics tells us that we could keep using mathematics even if we suddenly found out that mathematical objects don't exist. As Yablo puts it:

Mathematics does not lose its point either if the *mathematical* realm disappears – or, indeed, if it turns out that that realm was empty all along. [Yablo 2005, p.88 (emphasis in original)]

In fact, this is proved also by the mental experiment of the Oracle, which we discussed in section 3.3.3.

This important feature of mathematical discourse is what Yablo calls *the expressive indispensability* and amounts to the claim that mathematical discourse is indispensable, although mathematical objects are not. To be expressively indispensable means that with mathematical discourse we can express things we cannot express otherwise, and, when we say that mathematical objects are dispensable we are saying that for mathematical discourse to be useful we don't need mathematical objects to exist. This means that the expressive indispensability of number-talk says nothing about numbers. The latter can either exist or not, but this does not change the expressive role of mathematical discourse. In other words, saying that mathematical objects are dispensable means that the things we express via mathematical discourse have nothing to do with mathematics itself, but rather are about the concrete world. As Yablo puts it:

The economist need not believe in the average family to derive representational advantage from it (“the average family has 2.7 bank accounts”). The psychiatrist need not believe in libido or ego strength to derive representational advantage from them. Why should the physicist have to believe in numbers to access new contents by couching her theory in numerical terms? [Yablo 2005, p.94]

⁶⁰ See Balaguer(2013) for the details of the argument.

Also physicians take important advantages from the use of mathematical discourse, in the same way as economists and psychiatrists do in Yablo's example. As a matter of fact, it is convenient for physicians to use mathematical terms because without the reference to numbers and other mathematical objects they will face severe expressive problems, namely, they should use "uncountably many sentences, employing an uncountable number of numerical adjectives."⁶¹ Numbers, thus, turn out to be *representational aids* because they allow the indirect expression of facts not concerning numbers that we cannot directly express due to a lack of expressive resources. I have talked about numbers over and over again throughout the chapter, but I should explain one more time this important role performed by numbers, through an example from Yablo's work:

They [i.e. numbers] function *as representational aids*. This is how butterflies function in "I had butterflies in my stomach" and numbers function in "the number of Martian moons is 2." [Yablo, 2001, p.81 (emphasis in original)]

Furthermore, Yablo adds, this should not come as surprise because Field himself in his works defines truth in a deflationist way, that is, as a merely expressive device that does not correspond to any substantive property. Field doesn't need to take truth in an ontologically serious way in order to use the fragment of discourse involving it and, similarly, according to Yablo we don't need to take numbers and other mathematical objects in an ontologically serious way in order to use mathematical discourse. In both cases we are talking about representational aids that help us in the indirect expression of things not concerning respectively truth and numbers, that is, things about the concrete world that we cannot express in any direct way.

3.4.2 Fictionalism about Truth-Talk

Remember very briefly the aim of Armour-Garb and Woodbridge's work:

To belabour the point, what we will show is that if Yablo's considerations in favor of some sort of fictionalism about mathematics compel (and, for what follows, we will assume that they do), then we have a reason for thinking that T-deflationists ought to endorse or adopt some sort of fictionalism about truth-talk. [Armour-Garb and Woodbridge 2014, pp.93-94]

Therefore, the purpose of the two philosophers is not the development of a specific fictionalist theory about truth-talk, but rather to argue in favour of two things: on the one hand, they want to support the idea that there is a genuine compatibility between fictionalism about truth-talk and deflationism; and, on the other hand, they maintain there are several advantages for a deflationist in assuming a fictionalist perspective of truth-talk. What, in particular, this

⁶¹ Yablo[2005], p.94.

fictionalist perspective is, the authors don't tell us in the paper we are referring to.⁶²

The starting point for Armour-Garb and Woodbridge's argument is the fact that Yablo himself refers to deflationism in his own argument, and this, according to the two philosophers, is what legitimates the analogy between mathematics-talk and truth-talk. In fact, as we saw, Yablo himself refers to the analogy with deflationism in order to support his own thesis. He claimed that Field, as a deflationist, defines truth merely as an expressive device and, for this reason there is no need to take truth in an ontologically serious way in order to use the discourse about truth. Yablo makes the same point for numbers: we can be nominalists and we can take numbers in an ontologically non-serious way without being forced to give up the discourse about numbers. In other words, if truth is dispensable and the discourse about truth is expressively indispensable – that is, truth-talk can still be used even if the word “true” and its cognates refer to no substantive property at all –, then we can make a similar point about numbers as well: numbers don't exist and this is why they are dispensable, but at the same time the discourse about numbers is expressively indispensable, although number-terms pick out no object at all.

At this point, Armour-Garb and Woodbridge claim, it is possible to continue the analogy between deflationism about truth talk and nominalism in philosophy of mathematics by reversing it. In other words, according to Yablo, we can maintain that everything we can say of numbers and number-talk can be also said of truth and truth-talk. In particular, if Yablo is right when he says that nominalists should assume a fictionalist perspective of number-talk, then it is perfectly legit to claim that the same holds also for truth, namely, that deflationists should assume a fictionalists perspective of truth-theoretic discourse.

The argument advanced by the two philosophers goes on by showing that every step of Yablo's argument in favour of mathematical fictionalism can be reformulated in order to work for truth-theoretic fictionalism as well.

First, Armour-Garb and Woodbridge focus on the expressive role of the two regions of discourse. Yablo, as we saw, claims that the role of numbers is to function as representational aids, because they help us in the expression of “infinitely many facts in finite compass”⁶³ and those facts have nothing to do with numbers. In other words, numbers help us in the expression of facts not about numbers that we cannot express otherwise. Moreover, this role performed by numbers does not presuppose the existence of numbers themselves. However, we can make the same argument about truth-talk. Indeed, the role performed by truth is to function as a representation aid in the indirect expression of infinite conjunctions and disjunctions that we cannot express otherwise, and the real content of this conjunctions and disjunctions has nothing to do with truth itself. Hence, truth helps us in the indirect expression of facts not about truth that cannot be expressed otherwise.

⁶² As we will see at the end of the chapter (section 3.4.3), in Armour-Garb and Woodbridge[2015a], the two philosophers also propose a specific and well structured fictionalist theory of truth-talk. But, for the time being, let's focus on Armour-Garb and Woodbridge[2014], in which an approach of this kind has not been provided yet.

⁶³ Armour-Garb and Woodbridge[2005], p.101.

Moreover, this expressive role performed by the truth predicate does not require the existence of a substantive property of truth.

The second peculiar feature of Yablo's account is the introduction of the element of fiction. If the use we make of number-talk is limited to the expressive role of numbers that corresponds to the idea that numbers are representational aids – namely, that number-talk is not directly about numbers as it might appear, but rather it indirectly expresses things about the concrete world that we cannot directly express because of an expressive lack of our language–, then according to Yablo we must recognize that number-talk can afford this task via the operation of pretence, that is, via certain elements of fiction. As Armour-Garb and Woodbridge emphasize:

Functioning as a representational aid in this way – making as if to represent one thing for the purposes of representing something else, indirectly – just *is* to operate *via* some element of fiction. [Armour-Garb and Woodbridge 2005, p.105 (emphasis in original)]

Hence, the conclusion is that if we assume the nominalistic thesis that numbers don't exist, we obtain, as a result, that number-talk involves certain elements of fiction in order to perform its main role. Similarly, we can use the same argument for the discourse about truth. According to deflationism, the use we make of truth-talk is limited to its expressive role (in the same way as the use we make of number-talk is limited to its expressive role). Moreover, as we just saw, also from a deflationary point of view, truth functions as a representational aid, meaning that truth-talk is not about truth, but indirectly expresses things not about truth. Then, by miming Yablo's argument, we can say that since truth-talk functions as a representational aid, then it works through elements of fiction. Therefore, the conclusion is that if we assume the deflationist theses, then we obtain that truth-talk involved elements of fiction in order to perform his main role.

To reject this argument, the advocates of the opposite theories (realists about number-talk and inflationists about truth-talk) must provide at least one example where numbers and truth play an essential role in the expression of facts that really involve numbers and truth. In the case of truth, inflationists has the burden of prove that there are cases where “true” does not merely function as an expressive device. One could think that blind ascriptions might represent a problem in this respect, meaning that they seem to involve an instance of “true” that serves in the expression of facts that really involve the property of truth. Actually, as we already saw in section 2.1.4, blind ascriptions represent a paradigmatic example of the utility of the truth predicate as representational aid, because it helps us in the indirect expression of what we cannot directly express given our finitude. This is exactly what, by miming Yablo's argument, allows us to characterize truth-talk as fictional. Hence, Armour-Garb and Woodbridge conclude:

T-deflationists should conclude that the functioning of truth-talk involves an element of fiction. Put differently, if Yablo is right about number-talk, then T-deflationists should understand truth-talk to function through the operation of a fiction. [Armour-Garb and Woodbridge 2014, pp.109-110]

We can better explain the issue through a useful example. To say that the expressive role of truth is similar to the role of numbers means saying that the sentence “the Martian moons are 2” of numbers-talk is very similar to the sentence “what Bob said yesterday is true” of truth-talk. In the former case, the word “2” does refer to nothing because numbers don’t exist, and, as we saw, it functions as a representational aid in the indirect expression of facts not about numbers, that is, facts about the real world or, in the specific case, about Mars and its moons. Similarly, in the latter case, “is true” refers to no substantive property of truth, but rather it functions as a representational aid in the indirect expression of facts not about truth, namely, what Bob said yesterday. This feature is what, according to Yablo, allows us to talk about fiction, because, in particular, allows us to claim that numbers-talk – and, derivatively, truth-talk – works through elements of fiction.

3.4.3 More on Fictionalism about Truth

The point of the argument developed by Armour-Garb and Woodbridge we just saw is to prove that a deflationist should embrace a fictionalist perspective of truth-talk or, more correctly, that this should be the case provided that Yablo is right about mathematical fictionalism. What the authors do here is limited to this purpose, that is, to motivate the position of truth-theoretic fictionalism in general, and they don’t go into details by specifying how this specific account must be structured.

As we stressed over and over, the works on this matter available in literature are extremely limited. In particular, two are the fictionalist accounts about truth-talk we could find out. Let’s now briefly sum up their main features.

The first and less structured account is the one developed by Alexis G. Burgess in his PhD dissertation and then briefly mentioned in a book he recently published with his father, John P. Burgess.⁶⁴ To be central in this account is the distinction between what one says *within* the fiction and what one says *about* the fiction, that is the same distinction we talked about in section 3.1.6. Remember that it is possible to understand these two different kinds of discourse in the following way: what one says within the fiction is the intra-fictional discourse that occurs among non-specialist people in the middle of a street; on the contrary, what one says about the fiction is the extra-fictional discourse that occurs in a seminar room among philosophers. By using the same examples of Burgess and Burgess, the pretense is required by Tarskian biconditionals that can be represented via the following sentence:

- (14) To assert that some proposition is true is just asserting that proposition.

Hence, (14) is an example of what one says *within* the fiction. On the contrary, what one says *about* the fiction can be something like the following sentence:

⁶⁴ Burgess and Burgess[2011].

- (15) The liar paradox shows that the notion of truth is ultimately incoherent.

The difference between (14) and (15) lies in the fact that while the attitude of the speakers towards (15) is of genuine *belief* – meaning that who speaks really believes that the notion of truth is incoherent and that this is highlighted by the liar paradox –, the attitude towards (14) is merely of *pretense of belief*. What is required for someone to really believe a sentence like (14) is to interpret it fictionally, that is, to read it as being part of a fiction. The relevant fiction must be specified via the addition of an appropriate fictional prefix to the sentence, i.e. “according to the fiction of truth.” Hence, things are very similar to those about the following sentences:

- (16) Sherlock Holmes is a detective.

and

- (17) Sherlock Holmes is a character invented by Sir. Arthur Conan Doyle.

Sentence (15) is similar to sentence (17) because in both cases we really believe their literal contents, or, in other words, in both cases the real content exactly correspond to the literal. In the case of (17), we really believe that Sherlock Holmes is an invented character and that its author is Conan Doyle. Instead, (14) is very similar to (16) because in both cases we don’t really believe their literal contents, but we only pretend to believe it. If we consider (16), as well as (14), we realize that we merely have the attitude of pretence of belief towards it. We only believe the real content of these sentences, which is different from their literal one.

The second truth-theoretic fictionalist account I’m going to mention here is more recent and structured than the previous one. I’m talking about the *SPIF account* – namely, *Semantic Pretence-Involving Fictionalism* – developed by Armour-Garb and Woodbridge in their last book.⁶⁵ The point of their analysis is to provide a solution to what they call “philosophical puzzles or problems”⁶⁶ that are nothing but semantic paradoxes, the most important of which is the liar.

If in the previous section we analyse the argument suggested by the two philosophers in order to motivate the idea that a deflationist should be fictionalist about truth-theoretic discourse, here we are going to take into account the specific fictionalist theory developed by Armour-Garb and Woodbridge. As we mentioned in section 3.2.3, the two authors develop their own fictionalist theory on the basis of the ideas of Kendall Walton about fiction and its constitutive elements. Following this line, truth-talk, according to Armour-Garb and Woodbridge, is a game of make-believe such that it allows the indirect expression of certain things we couldn’t directly express. As we know, however, Waltonian games of make-believe have very specific features, namely, they are made of the following components: the *props*, the stipulated fictions, and the principles of generation. The great credit of Armour-Garb and

⁶⁵ Armour-Garb and Woodbridge[2015a].

⁶⁶ Ibid., p.ix.

Woodbridge with this account is of having specified what are the proper *props*, the stipulated fictions of truth-talk and the principles of generation that, starting from the other two elements, allow the functioning of the game of make-believe. The *props* – that, remember are elements of the real world that function as supports for the game of make-believe – of truth-talk are all those expressions such as “is true”, “is not true”, “is false”, “truth”, “falsity”, their cognates and the “that”-clauses that are linked to those expressions in propositional-talk. The stipulated fictions about these *props* are the following three:

- i. “is true”, “is false” and their cognates describe objects as having or lacking certain properties.
- ii. “truth” refers to the property attributed with the expression “is true”, and in a similar way for “falsity”.
- iii. “is true” is predicated of propositions. Other kind of objects can have the properties that “is true”, “is false and so forth attribute only derivatively, in virtue of expressing a proposition that has the relevant property.⁶⁷

The first stipulated fiction, i.e. i., tells us that “is true”, “is false” and so forth function as genuine predicates only within the fiction. Moreover, it’s worth noting that the second stipulated fiction holds that within the fiction of truth the truth predicate (and the falsity predicate as well) corresponds to a property that takes the name of “truth” (“falsity”). Finally, iii. specifies which objects the truth predicate is predicated of within the game of make-believe, i.e. propositions.

Lastly, Armour-Garb and Woodbridge provide the game of make-believe with the principles of generation on the basis of the *props* and the stipulated fictions. Through these principles of generation we can construct new fictions within the game of make-believe. The principles of generation for truth-talk are four:

- iv. The pretences involved in an utterance of “that p is true” are prescribed iff p.
- v. The pretences involved in an utterance of “that p is false” are prescribed iff $\neg p$.
- vi. If S_1 and S_2 are sentences that are alike except that one has a subsentence “p” and the other has a subsentence “p is true”, then one can infer S_1 from S_2 , and vice versa.
- vii. If S_1 and S_2 are sentences that are alike except that one has a subsentence “ $\neg p$ ” and the other has a subsentence “p is false”, then one can infer S_1 from S_2 , and vice versa.^{68 69}

The principles of generation are what allow the truth predicate to perform its essential role of expressive device. Thanks to the principles of generation the real content of the sentences of truth-talk arises, or, to be more specific, this is

⁶⁷ Ibid., p.131.

⁶⁸ Of course, vi. and vii. hold only in transparent contexts.

⁶⁹ Armour-Garb and Woodbridge[2015a], p.131.

so thanks to iv. and v. Hence, the first two principles of generation hold that when we utter a sentence like this:

(18) (The proposition) that snow is white is true.

what we indirectly mean via the truth predicate, i.e. its real content, is simply “snow is white.” The same goes for the following sentence:

(19) (The proposition) that snow is green is false.

Its real content, that is, what we indirectly want to convey through the use of the truth-talk (in this case of the falsity predicate) is nothing but “it is not the case that snow is green.” To restate the issue by using the vocabulary of Armour-Garb and Woodbridge, “that snow is white” operates as a designation expression that is content-connected to the proposition that snow is white. In other words, this means that the “that”-clause is a nominalization of the sentence “snow is white.” The same goes for falsity.

The last two principles of generalization, i.e. vi. and vii., guarantee the truth predicate to perform within the game of make-believe its expressive role that is universally known, namely, both to allow generalizations that cannot be expressed otherwise because of the infinite list of sentences that we should provide and to express blind ascriptions that we cannot otherwise express because of the impossibility to know the entire list of sentences constituting them. The former, remember, is the case of sentences like “everything Lisa says is true”, whereas the latter is the case of sentences like “what John said yesterday is true” when we don’t remember everything John said yesterday. In other words, vi. and vii. translate into the vocabulary of the game of make-believe a feature that the truth predicate has from an intuitive point of view, namely, the principle of intersubstitutivity.

Therefore, although apparently all these rules seem very complicated, they actually say nothing new about the truth predicate and its functioning. They simply adapt the features of the truth predicate we already know to the structure of the game of make-believe.

The choice to present the game of make-believe of truth-talk developed by Armour-Garb and Woodbridge along with all its main constitutive elements is due to the need of explaining and showing that we can, in fact, give a very clear and sharp structure of the functioning of the game of make-believe of truth-talk, that is, of fictionalism of truth-theoretic discourse. With this choice I wish to emphasize that the fictionalist account of truth-talk is not just a possibility for further developments for a deflationist, but rather it is a reality he can already endorse.

Chapter 4

Do Dialetheism and Deflationism Need Each Other?

4.1 Introducing the Issue

The present chapter starts out from an idea that can be briefly summarized with the following quote:

Dialetheism should be taken seriously by deflationists. Perhaps better: if we are right then deflationists should be dialetheists. [Armour-Garb and Beall 2003b, p.319]

Therefore, a deflationist should be a dialetheist. This is the idea Armour-Garb and Beall maintain in some of their works.¹ More specifically, the two philosophers in two different papers argue in favour of two theses: first, they aim to prove that deflationism is perfectly compatible with dialetheism; and, second, they claim this is not sufficient because deflationists not only can, but also should be dialetheists. The two theses I have just mentioned provide an affirmative answer to two questions corresponding to the titles of the two papers of Armour-Garb and Beall: “can deflationists be dialetheists?” and “should deflationists be dialetheists?”

Who adopts a theory about the nature of truth that doesn’t imply an ontological commitment towards the property of truth – i.e. deflationism – is naturally led to accept a dialetheic theory. According to Armour-Garb and Beall this is the only possibility available to deflationism in order to solve the liar paradox.

This chapter is inspired by this idea and my purpose here is to further investigate the issue. In particular, the chapter will be developed into two different directions.

First, I’ll try to provide an answer to the directly opposed question to the one presented by Armour-Garb and Beall. This means that in the first part of the chapter my aim will be to answer the question “assuming that deflationists should be dialetheists, will the opposite also be the case?”, or, in other more familiar words, “should dialetheists be deflationists?” In order to answer these questions, I’ll compare two different dialetheic theories: the theory developed by Graham Priest – with particular attention to the logic LP, that is the subject of sections 1.1.3 and following –, which doesn’t embrace a deflationary theory about the nature of truth;² and the theory more recently developed by J.C. Beall

¹ Armour-Garb and Beall[2001 and 2003b].

² Priest[1979 and 1987].

that is the perfect example of deflationary dialetheism.³ The cost-benefit analysis will allow me to provide an answer to the foregoing questions.

Secondly, I'll wonder whether things really are as Armour-Garb and Beall say, namely whether deflationists really should be dialetheists, or not. The answer will inevitably pass through the answer to another question: "why doesn't Hartry Field rely on a dialethic theory of truth in the development of a semantic for his deflationary theory and, on the contrary, prefers to develop his own and more complex paracomplete theory? In other words, why doesn't dialetheism suits him?" Therefore, also the second part of the chapter will be developed into a comparison between two theories: on the one hand, the deflationary and dialethic theory of Beall, and, on the other hand, the deflationary and paracomplete theory developed by Field.⁴ Similarly to the first part of the chapter, also in this case the cost-benefit analysis will help us to make an adequate answer to both the questions of this section.

The answer to all the questions in the chapter will be negative. This means that both there is no convenience for a dialetheist to be deflationist and also there is no convenience for a deflationist to be dialetheist, contrary to what Armour-Garb and Beall claim. The reason is not that the opposite is the case. In other words, the reason is not that is preferable for a dialetheist not to be a deflationist and for a deflationist not to be a dialetheist, but, instead, the actual reason is that dialetheists can be either deflationists or not equally, and the same goes for deflationists, they don't have any advantage in being dialetheists because being or not dialetheists takes the same advantages and disadvantages from a deflationary point of view.

Let's now focus on the arguments that lead Armour-Garb and Beall to the conclusion that a deflationist can and should be dialetheist.

4.1.1 Can Deflationists Be Dialetheists?

The first point of interest in Armour-Garb and Beall's paper that names the present section⁵ is to prove that deflationism is perfectly compatible with dialetheism, or, in other words, that there is nothing in the dialethic semantics and logic that prevent a deflationist from assuming it. In other words, as Armour-Garb and Beall put it:

Deflationists need not compromise their deflationary credentials by accepting dialetheism. [Armour-Garb and Beall 2001, p.594]

In particular, the proof provided by the two authors proceeds through the confutation of the arguments set forth by Keith Simmons,⁶ who claims that deflationism and dialetheism are incompatible.⁷

The argument by Simmons is based on the idea that central to deflationism's account are the following two theses:

³ Beall[2005 and 2009].

⁴ Field[2002 and 2008].

⁵ Armour-Garb and Beall[2001].

⁶ Simmons[1999].

⁷ Actually, both Simmons and Armour-Garb and Beall refer only to a specific type of deflationism, namely disquotationalism, but their arguments can be extended to cover also the other varieties of deflationism, and, for this reason, deflationism in general, taken as *genus*.

- (D₁) There's nothing more in the concept of truth than the principle of equivalence, that is, there's nothing more in the concept of truth than what is given by the disquotation of the quoted-name of a sentence.
- (D₂) The truth predicate is eliminable.

According to Simmons, the existence of ungrounded sentences⁸ is problematic for a deflationist in a twofold way. The two problems faced by deflationists are the *eliminability problem* and the *explanatory problem*. I'm going to focus here only on the latter, because the former is a problem common to all deflationists, whereas the latter is a serious challenge to the deflationist who wishes to assume a dialethic approach.⁹ Let's see, now, how Simmons' argument proceeds.

The *explanatory problem* raised by Simmons can be summarized as follows: if we embrace a deflationary perspective and, at the same time, we assume a dialethic theory, the result is that some ungrounded sentences are such that their truth can be explained only in terms of falsity, but this is a problem for a deflationist, according to which falsity is not an explanatory property at all.¹⁰ Simmons is obviously referring to ungrounded and paradoxical sentences such as the liar. As a matter of fact, according to the dialetheist, as we saw in section 1.4.4, the liar is both true and false. While for a dialetheist there is no problem in asserting that falsity explains the truth of the liar (and vice versa), this is not the case for a deflationist, according to which falsity (and truth as well) cannot explain anything at all. According to Simmons, this is the reason of the incompatibility between deflationism and dialetheism, namely, adopting a dialethic perspective will lead the deflationist to compromise his own theory about the nature of truth.

Armour-Garb and Beall's reply is based on the analysis of the dialethic theory. In fact, they claim, it is not true that according to the dialetheist falsity is invoked in the explanation of the truth of the liar and that truth is invoked in the explanation of its falsity, contrary to what appearances seem to suggest. If we take the two explanations together instead of taking them in isolation, we will see that they lead to a vicious circle, the conclusion of which is the rejection of both of them. As the two authors put the issue:

The falsity of l [i.e. the liar sentence] is implicated in the explanation of its truth, which, in turn, is implicated in the explanation of its falsity. Since the two explanations are symmetrical, if the dialetheist is correct in providing the one, she is correct in providing the other. But she cannot

⁸ Remember that a sentence is *ungrounded* if it refers neither to non-truth-theoretic facts nor to grounded sentences. In other words, a sentence is ungrounded if it contains at least one ineliminable occurrence of the truth predicate.

⁹ For the sake of completeness, it is useful to briefly explain what the *eliminability problem* says. Following D₂, the deflationist is committed to the eliminability of the truth predicate. However, the existence of ungrounded sentences, such as the liar sentence and the truth-teller, poses a serious problem for a deflationist, because such sentences involved ineliminable occurrences of the truth predicate. Therefore, as a consequence, we obtain that it is not possible to provide a deflationary analysis of ungrounded sentences.

¹⁰ The same goes for truth: the falsity of ungrounded sentences can be explained only in terms of their truth, but, according to deflationism, truth is not an explanatory property at all.

be correct in providing both. Hence, we conclude that she should provide neither. [Armour-Garb and Beall 2001, p.603]

Since neither deflationism nor dialetheism explain the truth of the liar through its falsity (and vice versa) and since this is the only argument produced so far that aims at proving the incompatibility between the two theories, Armour-Garb and Beall correctly conclude that there is nothing in the dialethic account that compromises the spirit of the deflationary theory of truth.

If the deflationist doesn't invoke truth or falsity in the explanation of respectively the falsity and the truth of the liar, then what's the explanation provided by the dialethic deflationist to this paradoxical sentence? According to Armour-Garb and Beall the explanation is provided by pointing out that both the assertion that the liar is true and the assertion that the liar is false directly follow from the deflationary theory, that is, they are logical consequences of the equivalence principle that is at the root of the theory, along with the identity between L and "L is false", and some other principles of logic.

To conclude, without the argument provided by Simmons, there is nothing allowing us to declare the incompatibility between deflationism and dialetheism, leading, in so doing, the way to a possible account that unifies the two approaches.

4.1.2 Should Deflationists Be Dialetheists?

In this section I'm going to put forward the next argument by Armour-Garb and Beall that will take the discussion one step forward. If in the previous section the two authors showed that deflationists can be dialetheists because there is nothing that prevents the union between the two theories, here, in this chapter, I'm going to show their next argument, according to which deflationists should be dialetheists, that is, there is an outright advantage for a deflationist in assuming a dialethic theory of truth.¹¹

As we might imagine, even in this case the key motivation lies in the liar. After all, this is not surprising because the liar is one of the main independent reasons that motivate the dialethic account.¹² In the opinion of many, this is all the more true for a deflationist, who seems to feel the blow of the liar more than others.¹³ However, according to Armour-Garb and Beall, there is a very simple and immediate way to solve the problem of the liar, namely, to adopt a dialethic theory of truth. Let's now see how the argument develops.

The real problem, the most urgent and of a difficult solution, that a deflationist faces is the so-called *delimitation problem*, that is, the problem of specifying the domain of a theory of truth, or, in other words, of specifying the proper truth-bearers of such a theory, i.e. the entities which "true" (and "false")

¹¹ Armour-Garb and Beall[2003b].

¹² Priest[1987].

¹³ See section 2.3.3 for an overview about the problem that the liar raises to deflationism and for the solutions that have been provided: the first attempt suggests to consider the problematic sentences, such as the liar, as expressing no proposition at all; the others, instead, are basically based on *ad hoc* restrictions of the T-schema. However, all of these proposals turn out to be inadequate for one reason or another.

can be predicated of. In particular, the problem raises the following issues. What are the proper truth-bearers according to a deflationist? Or, more specifically, are all the sentences proper truth-bearers or, instead, some of them must be ruled out because of the consequences they lead to? We are obviously referring to particular sentences, such as the liar. As Armour-Garb and Beall highlight, the theories that don't deflate truth are not subject to the delimitation problem:

Both the simple correspondence theory and Kripke's theory are able to delimit the proper truth bearers in such a way as to exclude the Liar. [...] Moreover, both rely on facts about the (substantive) property of truth – correspondence-theoretic and representational principles – in order to effect their delimitations. [Armour-Garb and Beall 2003b, p.309]

Hence, the theories rival to deflationism are not only capable of specifying the proper truth-bearers in such a way to avoid the consequences produced by the liar, but they are also able to do so through facts exclusively about the nature of truth.

Is the deflationist able to do the same? It seems he is not. As a matter of fact, the deflationist cannot, of course, appeal to a fact regarding the substantive property of truth because, as we stressed again and again, according to deflationism, there is no such substantive property. In fact, truth is a logical property and everything we can say about it is limited to the T-schema and its instances. The problem is that we cannot appeal to the T-schema in order to specify what are the proper truth-bearers, otherwise it would lead to a vicious circle, because which instances of the T-schema are to be admitted and which are not is exactly what we were supposed to figure out: in order to establish which sentences are suited to be genuine instances of the T-schema, we need to work on the basis of the T-schema itself, that is, we need to figure out which instances of the T-schema are problematic. Hence, the consequence is a vicious circle. The same problem has been highlighted in section 2.3.3. As we saw, some deflationary philosophers – e.g. Horwich –, who inevitably faced the problem due to the liar, proposed to restrict the class of correct instances of the equivalence principle, or, in other words, to restrict the class of proper truth-bearers. This means that not all sentences are appropriate to be truth-bearers, according to them. Which sentences are to be excluded? As we have already seen, there are two possibilities: we can rule out paradoxical sentences, or, alternatively, ungrounded sentences, in order to include also the truth-teller and its kin. The problem in both cases is that the solution turns out to be completely *ad hoc*. As we said, it is impossible to a priori know which sentences are paradoxical or ungrounded and which are not. We know some very famous ungrounded sentences – i.e. the liar and the truth-teller – but they are far from exhausting the entire list and there is no criterion based on the syntax of the sentences that can help us to know in advance which sentences are grounded or non-paradoxical and which are not. The only way is to look at the T-schema, but, as we just said, this is not an available solution because it leads to a vicious circle.

Therefore, the conclusion is that there is no *ad hoc* way to restrict the class of proper truth-bearers. As Armour-Garb and Beall puts the issue:

It appears that deflationists must allow that the range of application of “true” and “false” is maximally unrestricted – deflationary truth and falsity apply to any well-formed declarative sentence at all, whether grounded or ungrounded, paradoxical or not. [Armour-Garb and Beall 2003b, p.310]

That’s where dialetheism steps in the argument. If there is nothing – namely, no principle and no fact about the property of truth – that allows the restriction of the class of proper truth-bearers and if, for this reason, all the well-formed sentences are to be taken as proper truth-bearers,¹⁴ then it follows that the liar sentence is itself a proper truth-bearer. This means that the liar has a corresponding instance of the T-schema and this is a problem because, as we saw in section 1.4.2, this immediately leads to contradiction. As a consequence, we obtain that the deflationist is committed to at least one contradiction.

At this point the conclusion immediately follows from the argument. The possibilities we now have are two: either we reject deflationism, or we admit that the only available way to solve the problem is for the deflationist to embrace a dialethic theory. Obviously, the former is not an admissible option for a deflationist and, in any way, whoever decides to choose it, will do it on the basis that the only other way is unifying deflationism and dialetheism, but in this way she will reinforce Armour-Garb and Beall claim for which the only possibility for a deflationist is to be a dialetheist. Therefore, the only available alternative for a deflationist is to accept the contradiction that inevitably follows from her metaphysical intuitions about the nature of truth and from the principles that regulate its functioning. By so doing, the conclusion is that we have to admit that deflationism naturally leads to dialetheism.

4.2 Should Dialetheists Be Deflationists?

As we saw, the argument advanced by Armour-Garb and Beall is that not only deflationism is perfectly compatible with dialetheism, but also that deflationists should be dialetheists. The reason is that any other attempt to solve the problem common to all theories of truth, namely the liar, fails for one reason or another and, so, accepting the inconsistency of our language, through the admission of the existence of true contradictions, seems to be the only available way for a deflationist.

My aim in this section is to provide an answer to the complementary question of the one raised by Armour-Garb and Beall. If the purpose of the two American philosophers was to answer the question “should deflationists be dialetheists?”, my intention, here, is to answer the complementary question “assuming that deflationists should be dialetheists, will the opposite be the case? Should dialetheists be deflationists?” The very thesis I wish to support is that the answer to this question is negative: dialetheists shouldn’t be deflationists. However, the reason is not that dialetheists should be non-

¹⁴ As Armour-Garb and Beall[2003b] highlight, there obviously are some restrictions that we must make on the truth-bearers, but this constraints they have to obey are constraints of syntax and discipline, that don’t allow the deflationist to rule out problematic sentences at all.

deflationists, in the sense that deflationism is not an appropriate account about the nature of truth for a dialetheist, but rather that the dialethic logic and semantics fit well with every theory about the nature of truth. This means that dialetheists are not forced to choose a deflationary perspective in order to have the highest number of advantages, because deflationism has the same advantages of the others from a dialethic point of view.

It should be noted that Priest has already argued in favour of this neutrality of dialetheism with respect to the most important and known theories of the nature of true.¹⁵ The difference is that while Priest's aim can be summarized in the following way:

The question on the agenda is whether there is anything about the nature of truth that rules out dialetheism. [Priest 2000, p.306]

my purpose is different, because I wish to answer the following question: assuming that there is nothing about the nature of truth that rules out dialetheism and, so, assuming that dialetheism fits well with every account of the nature of truth, is there any account that is preferable over the others from a dialethic point of view? Or, more specifically, is deflationism such an account, in the sense that there is a convenience of a sort for the dialetheist in assuming a deflationary perspective instead of another that takes truth in a more substantive way?

The analysis will proceed via the comparison between two different dialethic theories: on the one hand, the theory developed by Graham Priest,¹⁶ who doesn't embrace a deflationary theory of the nature of truth; on the other hand, the theory more recently developed by J.C. Beall,¹⁷ that, provided with a transparent truth predicate, is the perfect example of a dialethic and deflationary theory of truth. The cost-benefit evaluations allows me to conclude that there is nothing in a dialethic theory that rules out a deflationary perspective of the nature of truth but, at the same time, there is nothing that prevents dialetheism from assuming a non-deflationary truth predicate.

4.2.1 Priest vs Beall: Dialethic Theories Compared

In order to answer the question "should dialetheists be deflationists?" we have to compare the two dialethic theories developed by Priest and Beall, so that we can set up a cost-benefit analysis. Furthermore, in order to see if the advantages of the dialethic and deflationary theory of Beall overcome those of the dialethic but non-deflationary theory of Priest, and if the disadvantages of the former are lower than those of the latter, we must figure out what are the main differences in the logical systems of the two theories and what they involve.

Before we examine the divergences at logical level, let's focus on the metaphysical issues. The first important difference between the two theories is philosophical. As we have stressed, both Priest's theory and Beall's are dialethic theories. Dialetheism, remember, was defined in section 1.2.2 as the

¹⁵ See Priest[2006], ch.2.

¹⁶ Priest[1987].

¹⁷ Beall[2009].

theory that there are true contradictions. The difference between Priest and Beall lies in the way each of them interprets the notion of “true contradictions”. In fact, while Priest recognizes the possibility of gluts in general, that is both in the semantic level – after the introduction of the truth predicate in the language – and in the true-free language, Beall maintains that gluts are only semantic, that is, they exist only at the semantic level. On the basis of this formulation, it seems a very banal difference, but it is actually substantial and very important. As Beall explains very well:

If one recognizes the possibility of gluts, in general, then the thought that some sentences are actually glutty is a natural course to follow. [Beall 2009, p.130]

Therefore, the trick is on the natural character of gluts. While according to Priest true contradictions are perfectly natural, according to Beall things are different. In fact, the latter maintains that gluts are not natural, but rather they appear as a consequence of our actions. In particular, gluts are the consequence of the introduction into our language of the device of (transparent) truth. Hence, gluts are, as he defines them, *spandrels* that we must accept as necessary if we want a truth predicate in our language. What does Beall mean when he says that gluts are spandrels? The explanation comes from Beall himself:

Spandrels of x are inevitable, and frequently unintended, by-products of introducing x into some environment. Originally, the term applied chiefly to architectural spandrels, those inevitable V-shaped areas that are by-products of arches. If you want arches in your design, you’re going to have spandrels. [...] Once spandrels enter the picture, one must decide what to do with them. One might ignore the spandrels; one might decorate the spandrels; one might try to hide the spandrels; one might do something else. Whatever one does, one cannot take them away, at least not without taking away the intended feature (e.g., arches) that brought them about. [Beall 2009, p.5]

Following and carrying on Beall’s metaphor, we can easily understand what spandrels of truth are in our language. The environment Beall talks about in this quote corresponds to the base language – that is, the language without the truth predicate – and the x – namely what has to be added to the base language – that in the architectonic metaphor corresponds to the arches, i.e. the feature we intentionally introduce into our environment, is the (transparent) truth predicate. The right conclusion is now easy to draw: the spandrels of truth, that is, the V-shaped areas of our language that are by-products of the truth predicate, are the problematic sentences we met throughout all the dissertation of which the liar paradox is a representative example. What can we do with these spandrels of truth? Beall suggests to accept them, that is, to simply admit that there are gluts in our language, or else we will have to remove the truth predicate from our language but, as we know from chapter 2, the truth predicate is an essential feature from the expressive point of view also for a dialetheist.

Therefore, the first important difference between Priest’s theory and Beall’s one can be sum up as follows: on the one hand, according to Beall, gluts

are merely semantic – that is, they involve only the semantic predicate of truth – and they are to be meant as the inevitable and unintended consequence of introducing the truth predicate into the language; on the other hand, according to Priest, gluts arise also outside of the semantic area and, so, they also appears into the base language, involving, in this way, also non-semantic predicates.

However, this philosophical difference between Priest and Beall’s views doesn’t imply any advantage for one of them, except that there seems to be a higher number of contradictions in Priest’s theory than in Beall’s one, because the former admit contradictions in more linguistic areas than the latter. As we will see, this is not a big deal because this will be swiftly balanced thanks to some logical feature of Beall’s theory that lack in Priest’s one, namely, the intersubstitutivity principle. Furthermore, the number of contradictions is not an adequate criterion on which to base our choice between the two theories. I’ll come back on this point at the end of the section, but for the moment let’s focus on the logical differences between the two dialethic accounts.

Since both theories are dialethic, it is obvious that both assume a paraconsistent logic and, thus, reject the rule of *ex contradictione quodlibet*, along with the disjunctive syllogism. The differences begin to arise when the truth predicate is introduced. In fact, let’s consider the formalization of the T-schema we already saw:

$$(T) \quad T((A)) \leftrightarrow A$$

Every dialetheist agrees that the conditional used in the formulation of the T-schema cannot be a contrapositional conditional. This means that the contraposition rule doesn’t hold for it. In particular, in a dialethic theory can hold neither weak contraposition:

$$(C-W) \quad A \rightarrow B \vDash \neg B \rightarrow \neg A$$

nor strong contraposition:

$$(C-S) \quad \vDash (A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$$

As a matter of fact, if contraposition held, the dialetheist would face very unpleasant consequences, because we would have that every sentence of the language would be true. In other words, if the conditional of the T-schema were a contrapositional conditional, the theory would be trivial, and, as we stressed, we want to avoid such a result. The argument to triviality is the following:

- | | | |
|----|------------------------------------|--|
| 1) | $\beta \wedge \neg\beta$ | [by assumption] |
| 2) | β | [by 1 and \wedge -Elimination] |
| 3) | $\alpha \rightarrow \beta$ | [by 2 and \rightarrow -Introduction] |
| 4) | $\neg\beta \rightarrow \neg\alpha$ | [by 3 and contraposition] |
| 5) | $\neg\beta$ | [by 1 and \wedge -Elimination] |
| 6) | $\neg\alpha$ | [by 4, 5 and <i>Modus Ponens</i>] |

Hence, either the dialetheist rejects one rule between contraposition and *modus ponens*, or from the assumption that there are true contradictions triviality follows. In fact, from the assumption that $\beta \wedge \neg\beta$ we obtain $\neg\alpha$ as a consequence, where α is whatever sentence of the language. Since *modus ponens* is a rule that nobody is prepared to waive because it represents very well our intuitions about the behaviour of the conditional in our natural language, the dialetheist is forced to reject contraposition in order to avoid triviality.

What implications does this move have? And why is it so important to our analysis? The most obvious and direct consequence of the rejection of contraposition is that if the conditional used in the T-schema doesn't contrapose, then in a dialethic theory we don't have the validity of the following schema:

$$(T-C) \quad \neg T(\langle A \rangle) \leftrightarrow \neg A$$

In particular, what the dialetheist rejects is the right to left side of (T-C), whereas he accepts the left to right one. In other words, we can eliminate untruth but we cannot introduce it from the negation of a sentence. As a result we have the lack for a dialetheist of the *full* transparency of the truth predicate. The reason is that in every transparent context we can always substitute A with $T(\langle A \rangle)$, and vice versa, *salva veritate*, but it is never possible to substitute (in a transparent context) $\neg A$ with $\neg T(\langle A \rangle)$ – the vice versa is the case – *salva veritate*.

What, in particular, is rejected by the dialetheist because of the lack of contraposition, is the following biconditional:

$$T(\langle \neg A \rangle) \leftrightarrow \neg T(\langle A \rangle)$$

More specifically, the dialetheist only validates the right to left side of the biconditional and rejects the opposite side. In this way, she prevents the attribution of truth to the negation of a sentence from being equivalent to the negation of the attribution of truth to the sentence itself.

As we can imagine, this situation satisfies only Priest, but not Beall, because by preventing the contraposition of the conditional used in the T-schema we are also preventing the full transparency of the truth predicate. However, as we saw in section 2.1.2, the transparency of the truth predicate is an essential feature for a deflationist, because it is the only way she has to move her metaphysical intuition about the nature of truth to the logico-semantic level. In fact, if there is no substantive property corresponding to the truth predicate and if the only role truth has is being a device that allows us to express some kind of generalizations that we cannot express in the language otherwise, then it is natural to think that at logico-semantic level there should be a property establishing that the attribution of truth to a sentence is equivalent to the sentence itself and that they should be intersubstitutable in every transparent context, i.e. a context without context-sensitive terms. It is intuitive and perfectly logical to think that if “snow is white’ is true” doesn't tell us nothing but “snow is white”, then it must be possible to substitute the former with the latter, and vice versa, in every transparent context, *salva veritate*, that is, without that any change in their truth-values occurs.

In view of this considerations, it appears very clear that for a dialetheist who means to embrace a deflationary approach about the nature of truth – such as Beall is – it’s not sufficient to have the truth predicate the way it is, that is, a truth predicate that, on the one hand, validates both sides of the T-schema but that, on the other hand, is not transparent because of the lack of contraposition. What measures do we need to take to fix the problem? We can introduce the intersubstitutivity principle that, along with the full T-schema, is intended exactly to make the truth predicate of the language for which we introduced the principle fully transparent. This is exactly the move made by Beall.

To conclude this first section, the main difference at logico-semantic level between the theory developed by Priest and the one developed by Beall is that the latter validates the Intersubstitutivity Principle in order to have the full transparency of the truth predicate, whereas the former doesn’t have this principle in his theory. Let’s now see, in next sections, what consequences this choice has for the two logicians and what further differences for the two theories it entails.

4.2.2 Gluts and Gaps

The first difference that immediately strikes one is that in Beall’s theory, thanks to the transparency of the truth predicate, all gluts are also gaps. In fact, as Field noticed,¹⁸ it is very easy to show how this equivalence appears. First of all, remember which are respectively the glutty and gappy claims:

$$\begin{aligned} \text{(GLUT)} \quad & T(\langle A \rangle) \wedge T(\langle \neg A \rangle) \\ \text{(GAP)} \quad & \neg(T(\langle A \rangle) \vee T(\langle \neg A \rangle)) \end{aligned}$$

Let’s now see, how the argument from the former to the latter proceeds, taking for granted that our theory is provided with the intersubstitutivity principle:

$$\begin{aligned} (1) \quad & T(\langle A \rangle) \wedge T(\langle \neg A \rangle) && \text{[by assumption]} \\ (2) \quad & A \wedge \neg A && \text{[by 1 and Intersubstitutivity Principle]} \\ (3) \quad & \neg\neg A \wedge \neg A && \text{[by 2 and Principle of Double Negation]} \\ (4) \quad & T(\langle \neg\neg A \rangle) \wedge T(\langle \neg A \rangle) && \text{[by 3 and Intersubstitutivity Principle]} \\ (5) \quad & \neg T(\langle \neg A \rangle) \wedge \neg T(\langle A \rangle) && \text{[by 4 and validity of } T(\langle \neg A \rangle) \leftrightarrow \neg T(\langle A \rangle)\text{]} \\ (6) \quad & \neg(T(\langle A \rangle) \vee T(\langle \neg A \rangle)) && \text{[by 5 and De Morgan rule]} \end{aligned}$$

Therefore, if our theory admits the existence of gaps or gluts – such as dialetheism does – and if in this theory the intersubstitutivity principle holds, then as a result we have that all gluts are gaps, and vice versa. The simplest way to show this equivalence is the following. By the intersubstitutivity principle – along with De Morgan rule applied to (GAP), that is a rule we can in no way reject if we wish to keep all of our intuitions about the functioning of the

¹⁸ Field[2008], p.382.

inferences we daily make safe – (GLUT) and (GAP) are respectively equivalent to what follows:

$A \wedge \neg A$

and

$\neg A \wedge \neg\neg A$

By the principle of double negation, $\neg\neg A$ is equivalent to A and, hence, the two claims above are equivalent.

Therefore, the first difference between Beall's and Priest's theory is that in the former all gluts are also gaps, whereas this is not the case in the latter.

Is this a problem according to Beall? In other words, is this equivalence between gluts and gaps a problem for a dialethic theory? No, it isn't. In fact, as Field well explains:

The conclusion that some or all gluts are also not gluts is of course perfectly coherent in paraconsistent logic, even if difficult for those of a more classical bent to understand. [Field 2008, p.367]

Hence, since the dialethic spirit involves the acceptance of a sentence and the acceptance of its negation, in the same way, there will be no problem for a dialetheist to accept that a sentence is both glutty and not glutty, or, in other words, that a sentence is both glutty and gappy.

Despite Field, as we already said, agrees that there is no problem from a dialethic point view in accepting the equivalence between gaps and gluts, he also maintains that the equivalence is the reason why Priest doesn't provide his theory with the intersubstitutivity principle. In other words, according to Field, Priest doesn't want the equivalence between glutty and gappy sentences for no good reason, since he is a dialetheist. As Field puts it:

Priest, as I've said, wants to be able to say of a sentence A that it is both true and false without concluding from this that it is neither true nor false. [Field 2008, p.365]

However, Priest immediately disproves this claim. In fact, he maintains that the resulting equivalence between gaps and gluts is not the real reason why he rejects the intersubstitutivity principle. Priest is very clear and cannot be misunderstood when he says:

Note that if one has the full contraposed T-schema [or, alternatively the Intersubstitutivity Principle],¹⁹ then gluts are also gaps, in the sense that $T\langle A \rangle \wedge F\langle A \rangle (= T\langle \neg A \rangle)$ entails $\neg T\langle A \rangle \wedge \neg F\langle A \rangle$. Field says that this is my reason for rejecting it. It isn't. [Priest 2010, p.130]

It seems, therefore, that there is no advantage in adopting Beall's theory in place of Priest's one, and vice versa. If it seems that there is no advantage in the equivalence between gluts and gaps, it is also clear that there is no advantage also in not having such equivalence. Furthermore, it seems that both Priest and

¹⁹ Anything allowing the truth predicate to be transparent.

Beall accept this fact.²⁰ For all these reasons, not to have the equivalence between gluts and gaps is not a good reason to choose Priest’s theory over Beall’s one, because, as Priest himself claims, this is not a problem for a dialethic theory at all.

In next section we will see what the real reason for Priest to reject the intersubstitutivity principle is and we will also try to figure out whether this can be a valid reason for allowing us to choose his theory in place of Beall’s one, or not.

4.2.3 The Number of Gluts

Priest maintains that preventing all gluts from being also gaps is not the real reason that motivates his choice of rejecting the principle of intersubstitutivity. What’s, then, this real reason? Priest himself tells us in the prosecution of the last quote:

The reason [for rejecting the contraposed T-schema] is that accepting it would turn any contradiction, $A \wedge \neg A$, into one of the form $T\langle A \rangle \wedge \neg T\langle A \rangle$, and “contradictions should not be multiplied beyond necessity”. In general, if A is true and false, I see no reason to suppose that $T\langle A \rangle$ should be false as well as true. [Priest 2010, p.130]

A new difference between Priest’s and Beall’s theories due to the intersubstitutivity principle is emerging. Indeed, the intersubstitutivity principle has, as a consequence, the validity of the following full biconditional:

$$T(\langle \neg A \rangle) \leftrightarrow \neg T(\langle A \rangle)$$

As we said in section 4.2.1, the lack of a conditional that validates contraposition has, for the dialetheist, the result of the non-fully validity of the above biconditional. As a matter of fact, without contraposition we can have only one side of the biconditional – namely, the right to left side –, whereas the other – namely, the left to right side – is not valid. If we introduce in our language the intersubstitutivity principle, what we obtain is that this biconditional becomes fully valid, i.e. both sides of the biconditional are valid.

The non-validity of the left to right side of the above biconditional is perfectly in line with Priestian ideas about untruth.²¹ In fact, according to Priest, there is an asymmetry between falsity and untruth. Priest defines falsity as truth of negation, whereas untruth is defined as the negation of truth. Since truth of negation is not equivalent to negation of truth, falsity is not equivalent to untruth. However, intuitively these two notions have something in common. After all, both of them are opposed to truth. As Priest puts it:

Falsity is one of the notions antithetical to truth. There is another, which I will call untruth. A sentence, α , is untrue if it not true, $\neg T\alpha$. [Priest 1987, p.69]

²⁰ And also some non-dialetheists. See Field[2008].

²¹ See Priest[1987], ch.4.9.

Moreover, also truth and untruth, in the same way as truth and falsity, are, according to Priest, exhaustive and non-exclusive. On the one hand, truth and untruth are exhaustive because it is the case that $T(\langle A \rangle) \vee \neg T(\langle A \rangle)$, since it is an instance of the excluded middle that is valid in a dialethic theory. On the other hand, truth and untruth are non-exclusive since there are sentences that are both true and untrue. The latter is the case of the strengthened liar, that is, the sentence that says of itself that it is not true and whose formalization is the following:

$$(Q) \quad \neg T(\langle Q \rangle)$$

From a dialethic point of view (Q) is both true and untrue.

What's, then, the difference between falsity and untruth? According to the dialetheist, untruth is in some way a stronger opposition to truth than falsity. As Priest says:

Truth and untruth are, therefore, "more inconsistent" than truth and falsity. [Priest 1987, p.72]

The reason of this higher inconsistency between true and untrue is the following. From the strengthened liar (Q) follows that there is at least a sentence that is both true and untrue and that formalized is: $\exists x(T(\langle x \rangle) \wedge \neg T(\langle x \rangle))$. However, the rule of excluded middle, that is valid in a dialethic theory, tells us that every sentence is either true or untrue, that is, $\forall x(T(\langle x \rangle) \vee \neg T(\langle x \rangle))$. By De Morgan rules, the latter is equivalent to $\forall x \neg(T(\langle x \rangle) \wedge \neg T(\langle x \rangle))$, that, by the principles of the quantifiers and the principle of double negation, is in turn equivalent to $\neg \exists x(T(\langle x \rangle) \wedge \neg T(\langle x \rangle))$, that is exactly the negation of our first claim. Therefore, it turns out that truth and untruth are both exclusive and non-exclusive. The same doesn't go for truth and falsity. The reason is the following. From the standard liar, similarly to what happens for the strengthened liar, follows that true and false are not exclusive, that is, $\exists x(T(\langle x \rangle) \wedge F(\langle x \rangle))$. The symmetry breaks out when we try to say that truth and falsity are also exclusive, i.e. $\neg \exists x(T(\langle x \rangle) \wedge F(\langle x \rangle))$. In fact, since we don't have the exclusion principle for truth and falsity, we cannot assert that they are exclusive. Hence, the incompatibility between truth and untruth is stronger than that between truth and falsity and this is the reason why a contradiction of the form $T(\langle A \rangle) \wedge \neg T(\langle A \rangle)$ seems to be more inconsistent than a contradiction involving falsity, i.e. $T(\langle A \rangle) \wedge F(\langle A \rangle)$.

However, along with the introduction in our language of the intersubstitutivity principle, the equivalence between $T(\langle \neg A \rangle)$ and $\neg T(\langle A \rangle)$ is resettled. What are the consequences of this move? And, are these consequences sufficient to establish which one between Priest and Beall's theories is preferable?

The main consequence is an increase in the number of contradictions. In fact, if the attribution of truth to the negation of a sentence is now equivalent to the attribution of untruth to the sentence itself, then what we obtain is that all contradictions of the form:

$$T(\langle A \rangle) \wedge T(\langle \neg A \rangle)$$

called “internal contradictions” because negation is internal to the scope of the truth predicate, give rise to an equal amount of contradictions of the form:

$$T(\langle A \rangle) \wedge \neg T(\langle A \rangle)$$

called “external contradictions” because negation is external to the scope of the truth predicate.

It’s easy to show why this happens. Since by the T-schema we have that $\neg A$ is equivalent to $T(\langle \neg A \rangle)$, and by the intersubstitutivity principle we have that $\neg A$ is equivalent to $\neg T(\langle A \rangle)$, it follows, by transitivity, that $T(\langle \neg A \rangle)$ is equivalent to $\neg T(\langle A \rangle)$. To make the equivalences obtained through the intersubstitutivity principle explicit, we can consider the following chain of biconditionals:

$$\neg A \leftrightarrow T(\langle \neg A \rangle) \leftrightarrow \neg T(\langle A \rangle)$$

from which it follows that all the contradictions of the form:

$$A \wedge \neg A$$

are equivalent to the following internal contradictions:

$$T(\langle A \rangle) \wedge T(\langle \neg A \rangle)$$

and to correspondent external contradictions:

$$T(\langle A \rangle) \wedge \neg T(\langle A \rangle)$$

The consequence is a multiplication of contradictions. Can this be a good reason for preferring Priest’s theory, that turns out to be more sparing in the number of true contradictions admitted, over Beall’s one? The answer is one more time negative. On the one hand, it is true that Priest’s theory admits only one contradiction of the form $T(\langle A \rangle) \wedge \neg T(\langle A \rangle)$ – namely, the strengthened liar – contrary to Beall’s theory which instead includes much more contradictions of that kind – i.e. all true contradictions of the form $A \wedge \neg A$ and $T(\langle A \rangle) \wedge T(\langle \neg A \rangle)$ are also contradictions of the form $T(\langle A \rangle) \wedge \neg T(\langle A \rangle)$ –, but, on the other hand, this is not sufficient to declare Priest’s theory the best dialethic theory. As a matter of fact, in this respect, there are two remarks that are imperative to be made.

The first observation that is worth noting is about the consequences that this multiplication of contradictions might have for a theory of truth. As also Beall himself highlights:

If, as on my position, truth is transparent (viz., our see-through device *truth*), then $\text{Tr}(\langle \neg \alpha \rangle)$ and $\neg \text{Tr}(\langle \alpha \rangle)$ are logically the same with respect to consequences. One ought not, as far as I can see, try to rid the virtues of *truth* in the name of Ockham-like concerns. [Beall 2009, p.131]

Since in Beall's theory $\text{Tr}(\langle \neg \alpha \rangle)$ is logically equivalent to $\neg \text{Tr}(\langle \alpha \rangle)$, the logical consequences of the former are also logical consequences of the latter, and vice versa. This means that admitting the equivalence between $\text{Tr}(\langle \neg \alpha \rangle)$ and $\neg \text{Tr}(\langle \alpha \rangle)$ imply no difference for the theory with regard to its logical consequences. Hence, it is not possible to maintain Priest's theory to be superior than Beall's one by appealing only to the parsimony on the number of dialetheias of the former. In fact, this would imply the use of an argument very similar to the Ockham's razor, but it would be of no help, because in a deflationary theory in which the truth predicate is transparent, such as Beall's one, admitting external contradictions, along with internal ones, doesn't compromise the theory if we look and focus only on its logical consequences.

Moreover, there is a second remark that is worth making here. Even if we don't take into account what we have already said about the logical consequences following from the two sentences $\text{Tr}(\langle \neg \alpha \rangle)$ and $\neg \text{Tr}(\langle \alpha \rangle)$, there is another problem that immediately appear to those who wish to establish which theory postulates the larger number of dialetheias. It is, in fact, impossible to make such a comparison on the number of true contradictions, because it is impossible – or, in any case, it is very difficult – to establish the real number of true contradictions a theory admits. If, on the one hand, it seems true that Beall admits more gluts than Priest because of the validity of the intersubstitutivity principle, on the other hand, it is also true, as we saw in section 4.2.1, that in Priest's theory gluts are admitted also in the true-free language, contrary to what accounted by Beall's dialetheism, where the so-called spandrels of truth are allowed only in the semantic level. The only way to establish which one has more contradictions in his theory is to figure out the exact number of gluts of both the theories, but this is not an easy operation to make because, as we already noticed in section 2.3.3, it is impossible to a priori know which sentences are paradoxical and which are not. In other words, we can't provide a full list of paradoxical sentences, because paradoxality does not depend from the syntax of the sentence and, for this reason, it is not possible to provide a syntactic criterion (nor a semantic one) that enables us to distinguish between “good” and “bad” sentences, where for “bad” sentences is meant those paradoxical sentences that spread gluts in the language.

It seems, thus, that at this level there is no possibility for us to determine which one between Priest and Beall's dialethic theory is the best and more adequate. Since these are the main differences between the two theories, it seems that it is not possible to establish which one is preferable if we look only at the logico-semantic level, because there is no substantial advantage or disadvantage for neither of them if we focus on the differences we just highlighted.

4.2.4 Contradictions Should Not Be Multiplied Beyond Necessity

If at logico-semantic level the differences between Priest's dialethic theory and Beall's one are not decisive at all, what do we have to rely on in order to establish whether dialetheists should be deflationists, or not, or, in other words, whether dialetheists should validate the intersubstitutivity

Do Dialetheism and Deflationism Need Each Other?

principle that is what allows the truth predicate to be transparent in a dialethic theory, or not?

Priest maintains that there is a legitimate and well-founded reason why he doesn't want to provide his theory with a transparent truth predicate, namely, because

Contradictions should not be multiplied beyond necessity. [Priest 1987, p.71 and p.116]

This very famous Priestian slogan summarizes a broader idea, namely that, according to a dialetheist, consistency is the case by default. This means that the concepts which we daily and commonly reason with are almost exclusively consistent. Inconsistency and dialetheias are extremely small in number. As Priest puts it:

The statistical frequency of dialetheias in normal discourse is low. Dialetheias appear to occur in a quite limited number of domains: certain logico-mathematical contexts, certain legal and dialectical contexts, and maybe a few others. Moreover, even in the domains where they do occur, very few contradictions are dialetheias. Hence most contradictions one normally comes across are not dialethic. [Priest 1987, p.116]

Therefore, not only most of the contexts are consistent and true contradictions belong only to a small number of domains of discourse, but also in these inconsistent contexts and domains, dialetheias are in a very limited number. As we said in section 1.2.2, one of the reasons that take Priest to claim in favour of the low probability of dialetheias is the common and successful use we do of quasi-valid rules. Using one more time Priest's words:

If dialetheias were common, we would expect quasi-valid inferences to go wrong quite frequently. But they do not. Hence they are not common. [Priest 1987, p.116]

We use quasi-valid rules successfully and this is why true contradictions are infrequent. From this fact it follows that it is usually rational to reject a contradiction, at least until we don't have grounds for believing that that contradiction is a dialetheia. For this reason, it seems correct to say that we don't have to postulate more contradictions than it is necessary, or, using Priest's words, we can restate the concept in this way: contradictions should not be multiplied beyond necessity.

Hence, according to Priest, that contradictions should not be multiplied beyond necessity is to be considered a valid and adequate criterion in order to evaluate the goodness of a theory of truth. Following his reasoning, since by the transparency of the truth predicate we have a real multiplication of the number of dialetheias, then it's better to reject the intersubstitutivity principle in order to keep the number of true contradictions limited. If, thus, this were a good criterion to evaluate a theory of truth, then Priest's theory would turn out to be preferable to Beall's deflationary theory. Hence, according to Priest, dialetheists should not be deflationists because otherwise they would multiply contradictions beyond necessity.

I agree with Priest when he says that contradictions should not be multiplied beyond necessity and I could also agree with him in saying that this is an adequate criterion on which to base our choice about the best dialethic theory. However, the problem is that this criterion cannot be taken unconditionally, so that we value a theory only on the basis of the number of contradictions it admits. Instead, a remark on how to intend the above criterion is required. In particular, it is indispensable to clarify what we mean with the word “necessity” used in Priestian slogan. In fact, it seems that different meanings of “necessity” lead to different answers to the question “what’s the right number of true contradictions a theory can admit?”

To this extent, the pressing issue now turns out to be the following: what does fall in the scope of necessity for a semantic theory of truth?

Contradictions should not be multiplied beyond necessity strictly depends on what necessity is. In fact, some people – i.e. Beall – maintain that having the principle of intersubstitutivity is a necessity for a theory of truth, and, according to this approach, it is clear that the multiplication of true contradictions that results from the principle of intersubstitutivity perfectly falls into the scope of necessity, and doesn’t transcend it, as Priest seems to suggest. At this stage someone could rightly wonder why according to Beall the intersubstitutivity principle is a *sine qua non* condition for a theory of truth. The answer to this question is fairly obvious. In fact, according to Beall, the intersubstitutivity principle is the only way to have the full transparency of the truth predicate for a dialetheist, because, as we saw, the dialetheist is forced to reject the contraposition of the conditional involved in the formulation of the T-schema, on pain of triviality. Hence, necessity, according to Beall, is having a theory that validates the intersubstitutivity principle, and the intersubstitutivity principle, in turn, is necessary in order to have a transparent predicate. If we keep going back in the regress, the next step is mandatory. We now wish to know why the transparency of the truth predicate is necessary for Beall. The answer is, one more time, very simple and intuitive. Beall supports a deflationary perspective about the nature of truth and the only way he has to transfer his intuitions about the nature of truth from the metaphysical to the logico-semantic level is via a transparent truth predicate, which allows him to substitute every sentence of the form “‘snow is white’ is true” with “snow is white”, and vice versa, *salva veritate*, in every non-opaque context. This is where the utility of the truth predicate lies from a deflationary point of view and, since there is nothing in the truth predicate but what this equivalence establishes, this is also everything the truth predicate consists in.

Now, what is necessity according to Priest? Contrary to what maintained by Beall, according to Priest, the principle of intersubstitutivity is not necessary at all in a theory of truth, because what is dispensable in Priest’s view is the transparency of the truth predicate. Why is not a transparent truth predicate necessary for Priest? The reason is that, in his metaphysical view, truth has not to be deflated. Priest doesn’t have a deflationary approach of the nature of truth, but rather a Dummettian one, meaning that he assumes the so-called teleological account of truth.²² According to him, truth is the *telos* of assertion,

²² Priest[1987], p.61.

Do Dialetheism and Deflationism Need Each Other?

that is, when we assert something what we aim at is truth, or, in other words, is saying something true. As Priest puts it:

The primary use of indicative sentences in a natural language is to make assertions. Asserting, like other human activities, has a *telos* or point, and the *telos* of asserting is truth. [Priest 1987, p.61]

From this point of view, it is pretty obvious thinking that truth is not a merely logical property; that is, everything we can say about truth is not limited to the T-schema, contrariwise to what deflationists claim. By using Priest's words again:

The inadequacy of the T-schema to characterise truth stems from the fact that it can capture only a certain set of logical relations between sentences. Yet the sentences of a language are part of a practice, and truth relates to how sentences are used in that practice, something that cannot be reduced to a set of logical relations between sentences. [Priest 1987, p.61]

Therefore, from a Dummetian perspective truth cannot be the notion described by deflationists because it cannot be reduced "to a set of logical relations between sentences", as Priest says, and this is exactly what deflationists do. The T-schema is not sufficient to grasp and describe everything we can say about truth. As a matter of fact, according to the teleological view, there is something more we can say about truth, namely, that is the *telos* of assertion. Following the Dummetian account of truth assumed by Priest, truth may not correspond to a substantive property – or, in other words, to a thick property –, in the way inflationists mean it, but, since it is what we aim at when we make an assertion, truth also cannot be a thin concept, in the way deflationists understand it. It is clear, then, that in this Dummetian view, the dialethic theory of Priest doesn't need a transparent truth predicate and, hence, it also doesn't need what makes the truth predicate transparent, i.e. the intersubstitutivity principle. For this reason, contradictions arising as a consequence of the introduction of the intersubstitutivity principle into the language are, according to Priest, totally beyond necessity.

To sum up, it is true what Priest says, that contradictions should not be multiplied beyond necessity, but, at the same time, it is also true that this depends from what necessity is, and what necessity is changes from theory to theory. On the one hand, according to Beall, necessity depends from the intersubstitutivity principle whose validity depends, in turn, from the necessity of a transparent truth predicate, which is a fundamental requirement for every deflationary theory of truth. On the other hand, according to Priest, necessity doesn't include the transparency of the truth predicate and, thus, the intersubstitutivity principle either, because the theory about the nature of truth assumed by Priest is not deflationary, but rather teleologico-Dummetian. Therefore, the conclusion seems to be that is true that contradictions should not be multiplied beyond necessity but what necessity is strictly depends from the underlying theory about the nature of truth.

4.2.5 From Semantics to Ontology

As we saw in the previous sections, we cannot establish which is the best dialethic theory of truth between Priest and Beall's one, or, at least, we cannot establish this if we look only at the logico-semantic structure of the two theories. The substantial difference between the theories is the intersubstitutivity principle, but it, by itself or along with the evaluation of its consequences, doesn't allow us to make a choice between Priest's and Beall's theory. Even the Priestian maxim according to which we don't have to multiply contradictions beyond necessity is of no help in this analysis, because what gets into the scope of necessity is different from theory to theory and primarily depends from the theory about the nature of truth that underlies the different logico-semantic accounts.

In this sense, the answer to the initial question – i.e. “should dialetheists be deflationists?” – is negative. The reason is not that it is more convenient for a dialetheist not to be deflationist, but rather that there is nothing from a logico-semantic point of view that pushes the dialetheist to choose a deflationary theory in place of another non-deflationary theory. In other words, from a logico-semantic point of view, she can equally assume either a deflationary perspective or a Dummetian approach of the nature of truth, without the consequences of this choice implying any advantage or disadvantage for the dialethic theory.

For this reason, it seems necessary to make a shift in the domain of the analysis. Since it's not possible to establish which is the best dialethic theory from a logico-semantic point of view, and since the only benchmark for the decision – namely, that contradictions should not be multiplied beyond necessity – strictly depends from what we take to be within the scope of necessity, and since what is necessary for a semantic theory depends from the theory about the nature of truth its advocates assume, then the metaphysical beliefs and the correspondent theories about the nature of truth will be the proper criterion at stake. Wondering whether it is better the dialethic theory of Priest or the one developed by Beall means nothing but wondering whether it is better the theory about the nature of truth subscribed by Priest or the one embraced by Beall. In other words, this means that to establish which is the preferable dialethic theory we cannot rely on the logico-semantic structure of the two theories because there is no fact of the matter as to which is the best dialethic theory from this point of view, but we should look at their metaphysical ideas and the corresponding theories of the nature of true they endorse.

To conclude, in order to answer the question “which is the best dialethic theory? Priest's or Beall's one?” we must shift from the logico-semantic to the ontological level, i.e. to the issue concerning the nature of truth. This means that in order to answer the previous question we have to answer the following: “is it better a deflationary or a Dummetian theory of truth?” No other analysis provides an adequate answer to the problem.

4.3 Should Deflationists Really Be Dialetheists?

In the previous sections, I argued in favour of the neutrality of dialetheism with regard to the theories about the nature of truth. In particular, I tried to prove that the two main dialethic theories – i.e. the Dummettian theory developed by Priest and the deflationary one developed by Beall – are equally good from the logico-semantic point of view. This result allows me to claim that deflationism is not the better theory about the nature of truth for a dialetheist, because there are other equally good theories, such as, for instance, the Dummettian theory that wants truth to be the *telos* of assertion.

Let's take a step back in the analysis and focus on the initial issue of the chapter, namely, "should deflationists be dialetheists?" As we know, Armour-Garb and Beall maintain that dialetheism is the best theory a deflationist could ever assume, because it doesn't face a whole series of problems which its rivals are subject to. What I'm going to do here is try to figure out whether Armour-Garb and Beall are right on this matter, or not, and whether there maybe is some other theory, different from dialetheism, that doesn't propose an *ad hoc* solution to the liar and that it's at least equally good as dialetheism is, or not. In other words, the analysis, from now on, will be aimed at showing if there is an alternative to dialetheism for a deflationary theory of the nature of truth.

To this extent, the question I wish to answer in this second part of the chapter is the following: should deflationists really be dialetheists? Similarly to the first part of the chapter, the analysis will proceed through the comparison between two theories. In this case, I'm going to compare two deflationary theories: on the one hand, the deflationary and dialethic theory developed by J.C. Beall²³ – that is one of the main characters we talked about in the first part of the chapter –;²⁴ and, on the other hand, the deflationary and paracomplete theory developed by Hartry Field.²⁵ Even in this case, the cost-benefit analysis will allow me to conclude that actually dialetheism is not the best possibility for a deflationist, contrary to what Armour-Garb and Beall maintain, because there is at least another theory that turns out to be equally good to dialetheism, at least if we limit our analysis to the logico-semantic structure of the two theories.

4.3.1 Field's Paracomplete Theory

In the same way as paraconsistent logics, paracomplete logics belong to the category known as non-classical logics. They are non-classical because they reject some peculiar laws or rules of classical logic. What laws or rules do paracomplete logicians reject? They reject the law of excluded middle. Hence, a

²³ Some aspects of the comparison will address Priest's theory instead of directly Beall's one, and they will refer to things said by Priest instead of Beall, but this is not a problem for the analysis because, as we already saw in the previous sections, the only substantial difference between Priest's and Beall's theory is the intersubstitutivity principle that holds only in the latter. Despite this aspect, hence, for a wider comparison between dialetheism and the paracomplete theory developed by Field, I'm legitimate to refer also to Priest's theory and works. The reason is the more completeness of Priest's works, contrary to Beall's ones, because they involve more aspects and allow us to make a broader analyses and comparison.

²⁴ Beall[2009].

²⁵ Field[2008].

logic is paracomplete if and only if rejects the classically valid law of excluded middle,²⁶ that can be formalized in the following way:

$$(LEM) \quad \vdash A \vee \neg A$$

Similarly to the notion of “paraconsistency” that, remember, literally means “beyond consistency”, the literal meaning of the word “paracompleteness” is “beyond completeness”.

The theory developed by Field is a paracomplete theory that satisfies the following desiderata:

- i. It’s a consistent theory. Hence, it doesn’t admit the possibility of gluts.
- ii. The language of the theory includes its own truth predicate.
- iii. It’s a generalization of the Kripkean theory of the fixed point and, at the same time, it solves Kripke’s problems about the conditional.²⁷
- iv. The truth predicate is transparent.
- v. Within the theory we can express the “defective” status of some particular sentences, such as the Liar.

In order to have these five desiderata, Field decides to adopt a paracomplete logic by rejecting the general validity of the law of excluded middle. Let’s now see the details of these features. First, i. assures us that Field’s theory is not also a dialethic theory, because it doesn’t admit the existence of true contradictions. Secondly, ii. tells us that the truth predicate used in the theory is defined in the language for which the predicate has been introduced, in the same way of Kripke’s theory and contrariwise to Tarski’s one, according to which there is a well-known difference between object language and metalanguage. Moreover, iii. specifies what the origin of Field’s paracomplete theory is, namely, it is a result of generalizing KFS, the internal theory of the fixed point. Field’s purpose is to generalize KFS in a way that also solves the problems due to the conditional. As a matter of fact, the conditional in KFS is not adequate, because it’s too weak to carry out ordinary reasoning.

²⁶ There are also other kinds of non-classical logics that reject the excluded middle, such as the intuitionistic logic. Contrary to the latter, however, Field’s paracomplete theory takes excluded middle to be valid for pure and applied mathematics. The only regions of the discourse where excluded middle doesn’t hold are those tied to truth and other similar concepts – that is, the area of the discourse concerning semantics – and those involving vague predicates.

²⁷ Actually, Field doesn’t generalize Kripke’s theory taken in its general meaning, but a very specific Kripkean theory. As a matter of fact, Field himself highlights the ambiguity in the use of the truth predicate made by Kripke. According to Field, when Kripke claims that he has defined truth for the language of his theory, he had improperly identified two different notions (one belonging to the metalanguage and the other to the object language): “having semantic value 1” and “being true”. There are different reasons why this identification is not profitable according to Field, like, for instance, the fact that it turns out to be implausible in the case of languages in which the quantifiers are unrestricted. Most important, the theory obtained from the separation of “having semantic value 1” and “being true” – called KFS – has some more favourable features than the theory which identifies the two notions – called FM. One advantage of KFS over FM that is worth noting is that it validates the intersubstitutivity principle. For the details see Field[2008], pp. 68-72.

More in details, the conditional is the material one that, remember, is defined in the following way: $A \rightarrow B =_{\text{def}} \neg A \vee B$. The problem is that the material conditional is totally inadequate in the case of a theory that doesn't validate excluded middle, such as KFS and Field's theory. For this reason, in KFS some very intuitive conditionals don't hold, such as $A \rightarrow A$, because, this would be equivalent to $A \vee \neg A$ that is nothing but an instance of the excluded middle, a non-valid rule in the theory. As Field puts it:

The lack of a conditional (and also of a biconditional) cripples ordinary reasoning. [Field 2008, p.73]

Therefore, it is necessary to provide the theory with a conditional that allows all the inferences we daily use, included $A \rightarrow A$. The complex construction made by Field has exactly this aim, namely to satisfy this requirement via the introduction of an adequate conditional by induction over the ordinals.²⁸ The fourth feature establishes that Field's theory must have the property of the full transparency of the truth predicate. In this way the intuitions about the nature of truth of the American philosopher – namely, that the truth predicate is nothing but a disquotational device – can be transferred from the metaphysical to the logico-semantic level. Similarly to Beall's theory, in order to obtain the full transparency of the language, in Field's theory both the full T-schema – i.e. both its side – and the intersubstitutivity principle hold. Finally, v. ensures that the theory is able to express the problematic – or, defective – status of some particular sentences, such as the liar.²⁹ In fact, in Field's paracomplete theory we cannot say that the liar is gappy, that is, neither true nor false. How can we satisfy the fifth requirement? Field proposes to define a determinacy operator within his theory.³⁰ This new operator is defined in the following way:

$$DA =_{\text{df}} A \wedge \neg(A \rightarrow \neg A)$$

or, alternatively, in the following way:

$$DA =_{\text{df}} (T \rightarrow A) \wedge A$$

where T is any logical truth.³¹

Field's idea is to characterize all the sentences of the hierarchy of liar paradoxes via the iteration of the determinacy operator, that is, via a stronger and stronger determinately truth. This determinacy operator is characterized in a standard way, since it obeys all standard rules of truth. In particular, it obeys

²⁸ More specifically, in order to obtain an adequate conditional, Field iterates the fixed-point construction through the ordinals. See Field[2008], ch.16.

²⁹ Also this problem is tied to Kripke's construction of the fixed point and, in particular, to KFS. In fact, in KFS is not possible to say something true about the liar. In other words, KFS is unable to express the "defective" status of the liar. The reason is that in KFS there is no possibility for defining a determinacy operator such that we can truly say that the liar is neither determinate true nor determinate false.

³⁰ The theory enables us to define a determinacy operator thanks to the new conditional.

³¹ Field also specifies that the determinacy operator applies to formulas that themselves contain the determinacy operator. In this way, he avoids the danger that the determinacy operator can bring a hierarchy of languages with it. See Field[2007], p.109.

Capture and Release in their rules form (respectively $A \vdash DA$ and $DA \vdash A$), but in their conditional form, it obeys only Release ($\vdash DA \rightarrow A$) and not Capture ($\nVdash A \rightarrow DA$).

With this idea in mind, Field can finally characterize the peculiar status of the liar sentence, (L), by saying that it is neither determinately true nor determinately untrue, and whose formalization is the following:

$$\neg DT(\langle L \rangle) \wedge \neg D\neg T(\langle L \rangle)$$

that, by the transparency of the language, is equivalent to the following sentence:

$$\neg DL \wedge \neg D\neg L$$

Hence, contrariwise to what happens in Kripke's account, in Field's theory we have a way to express the defective status of sentences like the liar, and this method involves the use of a determinacy operator.

The following question automatically arises: what about revenge? Can we construct a strengthened liar for Field's theory involving the determinacy operator? Or, using Beall's vocabulary, what happens with the inevitable spandrels of determinately truth? In fact, we can very easily formulate a sentence that says of itself that it is not determinately true, and that can be formalized as follows:

$$(W) \quad \neg DW$$

Of course, we cannot solve (W) by saying that is neither determinately true nor determinately untrue. The only possibility available to Field is to iterate the determinacy operator one more time so that we can say of (W) that is neither determinately determinately true nor determinately determinately untrue. The operation will be repeated for all the other liar sentences that inevitably arise because of this iteration. The result is that for every liar $(W_{\sigma-1})$ we can always say that it is $\neg D^\sigma W_{\sigma-1} \wedge \neg D^\sigma \neg W_{\sigma-1}$. How long can we repeat the iteration? The determinacy operator can be iterated into the transfinite, in a way that it blocks every new liar. Field introduces a hierarchy of levels for the notion of "determinately true" – or, in other words, he expresses the notion of "determinately true" through a hierarchy of operators \neg , such that we can say that a sentence is determinately true only at the next level.

This is the very general idea of the functioning of Field's theory. After having explained what are the requirements characterizing the theory, we can compare it with dialetheism in order to figure out which of the two theories is the best from a deflationist point of view.

4.3.2 Beall vs Field: Deflationary Theories Compared

In the light of what I said about the paracomplete theory developed by Field, I will now try to answer the fundamental question of this second part of the chapter: should deflationists really be dialetheists, as argued by Armour-

Garb and Beall? In order to do so, we must compare two different deflationist theories, individually developed by Beall and Field, so that we will carry out a cost-benefit analysis very similar to the one we did in the first part of the chapter. In order to see if the advantages of Beall's theory overcome those of Field's, and if the disadvantages of the former are lower than those of the latter, we must look at the main differences in the logical structure of the two theories and at the consequences they lead to.

What I intend to claim here is that, despite the two theories apparently differ in a very marked way – in fact, for instance, one is a paraconsistent theory that rejects the rule of *ex contradictione quodlibet*, whereas the other is a paracomplete theory that rejects the excluded middle –, they are not that far apart. What I aim to show is that the two theories are actually dual, especially if we focus our analysis on their logical structure. In particular, this will arise from the fact that the two theories turn out to have the same consequences, they face the same problems and the solutions they provide mirror each other. It's clear that if the two theories are dual, then a negative answer to the original question is inevitable because it's a natural consequence of this duality. As a matter of fact, if the two theories are dual with regard to the logico-semantic aspects, then it is natural to conclude that if they are good, they are equally good. This means that there's nothing in their logic that allows us to choose one over the other. Therefore, from this result follows that deflationists need not be dialetheists, contrary to what Armour-Garb and Beall claim, because there is no effective advantage for a deflationist in assuming dialetheism. He can also embrace an adequate paracomplete theory, such as Field's one. Hence, the deflationist has two equally good possibilities to choose from.

We can notice some base features that the two theories share and that help us to lay the foundation for the analysis I'm going to make in next sections.

First, we can show that the two theories – or, better, the logics assumed by the two theories, namely the paraconsistent and the paracomplete logics – can be traced back to the same starting point, namely the construction of the fixed point made by Kripke. The point here is not that the two theories actually developed in this way – and, in fact, Priest doesn't have in mind Kripke's construction in the development of his logic³², contrariwise to Field –, but rather what it's important is that both Priest's paraconsistent logic, namely LP, and Field's paracomplete theory can result from Kripke's construction of the fixed point. For this reason, we say that the two theories can have the same common origin.

Kripke's construction of the fixed point³³ is an account developed by the American philosopher with the aim of defining truth for a language containing the truth predicate. In this way, Kripke aimed at solving the well-known problem of Tarski's theory that requires a distinction between object language and metalanguage. Moreover, what Kripke's construction allows to do is to say that some peculiar sentences (the liar, the truth-teller, etc.) has an indeterminate truth-value, so that they are neither true nor false. This possibility is accounted for by allowing the truth predicate to be only partially defined. In order to obtain this result, we have to modify the way we interpret the truth predicate

³² Remember that also Beall's theory is based on LP.

³³ Kripke[1975].

and, in particular, we have to fix an extension and an anti-extension of that predicate.³⁴ The first step is to assign all sentences of the base language, i.e. the true-free language, to the extension or the anti-extension of the truth predicate, so that all sentences such as “snow is white” will belong to the extension and all the sentences of the form “snow is green” will belong to the anti-extension.³⁵ At next step, we will evaluate in the same way the sentences obtained from the attribution of the truth predicate to the sentences already evaluated and the new combinations that obtain a value accordingly.³⁶ In this way, we will have that the sentences such as “‘snow is white’ is true” will belong to the extension of the truth predicate and the sentences like “‘snow is green’ is true” will belong to its anti-extension. This procedure will be iterated until all sentences will be evaluated. In this way, we are going to have a partial interpretation of the truth predicate that becomes richer and richer as long as the evaluation is possible. Intuitively all attributions of truth to sentences that were in the extension will also be in the extension and all the attributions of truth to sentences belonging to the anti-extensions will also be in the anti-extension. This assures us that the truth predicate is transparent. Then, Kripke proves that there is a point where the interpretation of the truth predicate can no longer expand. In other words, as one proceeds through higher and higher levels, one eventually comes to a point where no new evaluation is possible. This is called by Kripke the *minimal fixed point*. All sentences in the minimal fixed point, that is, all sentences that have a truth-value in the minimal fixed point are called *grounded* sentences. The others are called *ungrounded*. The paradoxical sentences are those sentences that belong to no fixed point at all. Both the liar and the truth-teller are ungrounded sentences, because they are neither in the extension nor in the anti-extension of the truth predicate at the ground level (because they involve the truth predicate, and cannot be reduced to any sentence that doesn’t involve it) and they cannot be introduced through any subsequent evaluation step, hence, they also are not in the minimal fixed point. However, the liar, contrary to the truth teller, is also paradoxical because, while we can arbitrarily assign at ground level a truth-value to the latter, by getting in this way a fixed point including the truth-teller, the same doesn’t apply for the former. In fact, we cannot arbitrarily assign a truth-value to the liar at the ground level without entailing a contradiction.

The most interesting thing that is worth noting about Kripke’s construction of the fixed point is that it doesn’t correspond to a single theory,

³⁴ Where the extension and the anti-extension of a predicate, P, are defined in the following way:

- The *extension* of P is the set of objects which P is true of.
- The *anti-extension* of P is the set of objects which P is false of.
- P is undefined otherwise.

The third point guarantees the possibility for a sentence to be neither true nor false, when P is the truth predicate.

³⁵ This is a rough simplification. In fact, at first level we evaluate also sentences involving the truth predicate, where the possible value of the truth predication does not affect the evaluation. This is the case of disjunction in which one disjunct is true and the other can be anything – also a sentence involving the truth predicate –, or a conjunction in which a conjunct is false and the other can be anything – also a sentence involving the truth predicate –, etc.

³⁶ For instance, the sentence “ $2+2=4$ and $T(\langle p \rangle)$ ”, where T does not occur in p, is evaluated at second level because both conjuncts are evaluated in this level. On the contrary, the sentence “ $2+2=4$ and $T(T(\langle p \rangle))$ ” is not evaluated at second level because the second conjunct is not been assigned a truth-value yet.

but rather to a set of different theories that can be developed starting from Kripke’s account, depending on the interpretation we give to some notions of the construction and on the values we take as designated. In particular, the theories we can obtain from Kripke’s construction are four and correspond to the following:

- 1) *Classical gap theory*: in which the extension of the truth predicate is the set of sentences with value 1. In this case, we have the identification between “being true” and “having semantic value 1”.
- 2) *Paracomplete theory (KFS)*: in which the extension of the truth predicate is the set of sentences with value 1. Contrariwise to the previous, the present theory doesn’t identify “being true” with “having semantic value 1”.
- 3) *Classical glut theory*: in which the extension of the truth predicate is the set of all sentences with value other than 0. Like the first theory, here we have the identification between “being true” and “having semantic value 1”.
- 4) *Dialethic theory (LP)*: in which the extension of the truth predicate is the set of all sentences with value other than 0. Even in this case, likewise the second theory, we don’t have the identification of “being true” with “having semantic value 1”.

As we can intuitively understand from the above list, the identification of “being true” with “having semantic value 1” is what characterizes classical theories and what makes it different from non-classical ones. In fact, as we already saw in section 4.3.1, to split up the two notions imply the passage from a two-valued logic to a three-valued one.

At point 2 and 4 we can find our two non-classical theories: at point 2 is described KFS, on which Field bases his paracomplete theory; and at point 4 there is LP, the logic of paradox which every dialethic theory is based on.

To that extent, it is plausible to claim that the two theories can (at least in theory) have the same origin (even if, as we said but it is worth stressing, it is, in fact, not the case because Priest doesn’t refer to Kripke’s construction in the development of LP). This means that starting from Kripke’s construction of the fixed point it’s possible to develop both Field’s paracomplete theory and Priest and Beall’s dialethic ones.

The fact that the two theories can be developed from the same starting point proves and, at the same time, is proved from their truth tables. In fact, according to Kripke, the most appropriate schema for handling the connectives is Kleene’s strong three-valued logic (also known as K3). Their behaviour is explained in what follows:

- *Negation*. The negation of a sentence, A, is true (false) if A is false (true), and undefined if A is undefined.

A	$\neg A$
V	F
F	V

u	u
---	---

- *Conjunction.* A conjunction is true if both its conjuncts are true, it's false if at least one conjunct is false, and undefined otherwise.

A	B	$A \wedge B$
V	F	F
V	V	V
F	F	F
V	u	u
F	u	F
u	u	u

- *Disjunction.* A disjunction is true if at least one of the disjuncts is true, it's false if both the disjuncts are false, and undefined otherwise.

A	B	$A \vee B$
V	F	V
V	V	V
F	F	F
V	u	V
F	u	u
u	u	u

- *Conditional.* In K3 the conditional is the material one, so its truth table is constructed on the basis of the tables for negation and disjunction:
 $A \rightarrow B =_{\text{def}} \neg A \vee B$.

A	B	$A \rightarrow B$
V	F	F
V	V	V
F	V	V
F	F	V
V	u	u
F	u	V
u	F	u
u	V	V
u	u	u

Exactly as in the case of the truth tables for LP, the red part underlines the behaviour of the connectives when we introduced the new truth-value in the logic.

The first thing that we notice when we look at the truth table for K3 is that they seem to us very familiar. In fact, as we already pointed out in section 1.1.3, there is a further contact point between Field's theory and the dialetheic one: the matrix of their truth tables is exactly the same, namely that of K3. What changes is the interpretation they give to the third truth-value. In the case of dialetheism, that is, in LP, the third truth-value is "both true and false", whereas in the case of the paracomplete logic which Field's theory is developed on, i.e. KFS, the third value corresponds to "neither true nor false". However, the behaviour of the connectives with regard to those new values is the same. The difference can be explained using the notion of *designated values*. In the case of K3, and KFS as well, the only designated value is "true, whereas in the case of LP, the designated values are two, namely "true" and "both true and false". This can be shown from the division of Kripke's construction into the four theories we mentioned above. In fact, if we take into account the second and fourth theories, those corresponded respectively to KFS and LP, we immediately see what follows: on the one hand, the former is defined as the non-classical theory in which the extension of the truth predicate is the set of sentences with value 1, that means nothing but that the only designated value is 1; on the other hand, the latter is defined as the non-classical theory in which the extension of the truth predicate is the set of all sentences with value other than 0, that means nothing but that the designated values in LP are all values greater than 0.

Lastly, there are few final considerations that are worth mentioning, even if some of them are quite obvious at this point. First, I'd like to stress that both Field and Beall's theories validate not only all the Tarskian biconditionals – that is, the full T-schema –, but they also account for a fully intersubstitutable truth predicate. Therefore, as Beall himself underlines:

Our respective accounts both aim to be accounts of the same transparent truth predicate and its spandrels. [Beall 2009, p.95]

Secondly, the last remark is about the "degree of classicality" preserved by the two theories. As Beall marks, both Field's and Beall's language "enjoys a proper, classical fragment."³⁷ The size of this fragment is different, but it's remarkable and important to notice that both the logicians take mathematics to be fully classical.

To conclude this first section, a first glance to the logico-semantic structure of the two theories allows us to highlight that the differences are not such that they can justify the choice of one theory over the other. In the following sections we will look more in details at other less superficial aspects of the two theories.

³⁷ Beall[2009], p.95.

4.3.3 Curry's Paradox

The Curry's paradox belongs to the family of semantic paradoxes of self-reference. It takes its name from the American mathematician and logician that discovered it in 1942, namely, Haskell B. Curry.

The Curry's sentence is a sentence that says that if it is true than an absurdity follows. In other words, The Curry's sentence (C) says something like the following:

$$(C) \quad \text{True}(\langle C \rangle) \rightarrow \text{The Earth is flat}$$

Using the Diagonalization Lemma,³⁸ we can construct a truth-theoretic and formalized version of the Curry's sentence that can be stated in the following way:

$$(K) \quad T(\langle K \rangle) \rightarrow \perp$$

where \perp is any absurdity.³⁹ It's easy to show that from the Curry's sentence an absurdity follows. Let's consider the following argument:

(1)	$T(\langle K \rangle)$	[by assumption]
(2)	K	[by 1 and T-schema]
(3)	$T(\langle K \rangle) \rightarrow \perp$	[by 2 and construction of K]
(4)	\perp	[by 1, 3 and <i>Modus Ponens</i>]
(5)	$T(\langle K \rangle) \rightarrow \perp$	[by 1, 4 and \rightarrow -Introduction]
(6)	K	[by 5 and construction of K]
(7)	$T(\langle K \rangle)$	[by 6 and T-schema]
(8)	\perp	[by 5, 7 and <i>Modus Ponens</i>]

Hence, we have just proved that from the Curry sentence an absurdity obtains. The real problem is that \perp can be replaced by any other arbitrary sentence, but in so doing we will have that from the Curry's sentence every sentence follows. Therefore, the devastating consequence of the Curry's sentence is, one more time, trivialism.

What is worth noting is that the Curry's paradox is not a problem only for classical theories, but also for non-classical ones, included dialetheism and Field's paracomplete theory. In fact, as Beall says in the following quote, the Curry's paradox seems to be a bigger problem for non-classical theories than for classical ones:

Where Curry's paradox becomes especially significant is not with classical approaches but rather with certain non-classical approaches. In particular, Curry's paradox is a direct challenge to any non-classical approach that attempts to preserve one of the canvassed schemas – truth, sets, semantical properties – in unrestricted form. [Beall 2013]

³⁸ Cf. footnote 36 cap.1.

³⁹ \perp can also be a contradiction, or any other absurdity.

While classical theories of truth restrict one side or the other of the T-schema in order to solve the Curry's paradox,⁴⁰ this solution is available neither to dialetheism nor to Field's paracomplete theory, because each of them aims to keep the full T-schema safe. Moreover, the reason why non-classical theories suffer the Curry's paradox more than classical theories is that the absurdity following the Curry's sentence is not due to Boolean negation and, for this reason, it cannot be solved by non-classical theories in the same way they solve the other semantic paradoxes, e.g. the liar and its strengthened kin. As a matter of fact, the solution non-classical theories provide to the liar involve in some way a changing in the meaning of negation, because they incorporate one sort or another of non-Boolean negation. On one hand, as we saw, dialetheism embrace a paraconsistent logic, which is based on the fact that the inference from $A \wedge \neg A$ to an arbitrary B fails, since A can be a dialetheia, that is, since A can be both true and false. On the other hand, the paracomplete approach is based on the fact that the inference from $A \leftrightarrow \neg A$ to $A \wedge \neg A$ fails, since A can be neither true nor false. In both cases negation is not the Boolean one, which, instead, is also characterized by the validity of both the inferences they reject.

The main feature of the Curry's paradox lies exactly in the fact that it doesn't involve negation at all. In this way it prevents both the paracomplete theory developed by Field and dialetheism from blocking the derivation of the problematic conclusion that leads to trivialism via the same solution they provide for the liar. Hence, turning Boolean into non-Boolean negation is not the right way to go in order to solve the Curry's paradox.

What are, then, the solutions put forth by the dialetheist and the paracompletist? If we consider the argument of the Curry's paradox stated above, we can noticed that in order to block it we must reject one of the rules involved in it. In other words, we must reject at least one of the following rules: the full T-schema, *modus ponens* or the introduction of the conditional. Of course, neither Field nor Beall are willing to abandon the full T-schema because, as we already saw, one of the fundamental features a deflationary theory of truth must have is the transparency of the truth predicate, and the T-schema is fundamental for its achievement. Furthermore, also in Priest's account we have the validity of the full T-schema – that is, both its sides but not the contraposed T-schema – and, hence, also according to Priest, who doesn't have a deflationary perspective of the nature of truth, the T-schema cannot be rejected in order to block the derivation of the Curry's paradox. More in general, the T-schema is a very intuitive desideratum for every theory of truth, be it deflationary or not. The possibilities, thus, are only two: either we can waive *modus ponens* or we can abandon the rule of the introduction of the conditional.

⁴⁰ The validity of the full T-schema is exactly what classical logicians reject. In particular, it is questioned by those classical logicians who admit the existence of gaps or gluts (see section 4.3.3 for an overview on this matter). More specifically, on the one hand, classical gap theorists block the argument at step 7, by rejecting the right to left side of the T-schema, that is, by rejecting that the truth of a sentence follows from the sentence itself. On the other hand, classical glut theorists block the argument of the Curry's paradox at step 2, where the opposite side of the T-schema is involved, that means they reject the left to right side of the T-schema, that is, they reject that from the attribution of truth to a sentence the sentence itself follows.

The solution to the Curry's paradox proposed by Field is to block the argument at step 5, by rejecting the rule of the introduction of the conditional.⁴¹ Field maintains that the metalinguistic rule of the introduction of the conditional, that can be formalized as follows:

$$(\rightarrow\text{-I}) \quad \text{if } \Gamma, A \models B, \text{ then } \Gamma \models A \rightarrow B$$

is not validity preserving. According to Field, in order for the rule to preserve validity, an assumption must be added, so that we can have what follows:

$$(\rightarrow\text{-I}_R) \quad \text{if } \Gamma, A \models B \text{ and } \models A \vee \neg A, \text{ then } \Gamma \models A \rightarrow B$$

The reason is very intuitive and clear: we want the rule to be valid in all cases in which excluded middle is valid as well, that is, in all cases in which we are sure there are no problematic and defective sentences, that, on the contrary, require a revision of classical logic through the restriction of some classical rules.

There is, of course, an independent reason why Field restrict the validity of the introduction of the conditional to all cases in which excluded middle holds. For if that were not the case, his restriction would be totally *ad hoc* and its legitimacy would be call into question. According to Field, we cannot legitimately pass from the premise $\Gamma, A \models B$ to the conclusion $\Gamma \models A \rightarrow B$, because this is not the right inference. The proper inference allowed by the premise is the following: $\Gamma, A \models A \rightarrow B$. Moreover, one can assert a conditional from the assumption that the antecedent is false. Hence, we also have the following: $\Gamma, \neg A \models A \rightarrow B$. Therefore, since paracomplete logic allows the reasoning by cases, the result is that from the original premise follows $\Gamma, A \vee \neg A \models A \rightarrow B$, which, as a consequence, gives us the rule of the introduction of the conditional in its revised form, i.e. $(\rightarrow\text{-I}_R)$.

Let's now see what the solution to the Curry's paradox provided by Priest is. As we expect, the move made by Priest in LP mirrors that of Field. In LP, as we know, the conditional is the material one and, for this reason, Priest decides to block the argument at step 4 or 8, or both of them, by rejecting the *modus ponens*. *Modus ponens* is a rule also known with the name of elimination of the conditional, because it allows the elimination of the conditional, that is, enables us to move from a conditional to its consequent provided the assumption of its antecedent. Moreover, similarly to what Field does, Priest doesn't totally reject *modus ponens*, but, instead, he includes it into the set of quasi-valid rules,⁴² which, remember, are rules valid in all consistent contexts, but not in dialethic ones.

Also in Priest's case, the move is not *ad hoc*. As a matter of fact, the rejection of the *modus ponens* has an independent motivation that is not limited to the solution of the Curry's paradox and that is tied to the material conditional. As we saw in section 1.2.1, if the conditional is the material one, *modus ponens* is equivalent to the disjunctive syllogism that, as we know, is rejected – or better, it is restricted – by Priest in order to block the Explosion's argument.

⁴¹ Field[2008], p.281-284.

⁴² See section 1.2.2.

To sum up the section, the Curry's paradox turns out to be a problem for both theories because it involves negation neither in its formulation nor in its argument. Hence, the traditional solution the two theories provide for semantic paradoxes is not available for the Curry's. Both paracompletists and dialetheists are, thus, forced to find an alternative solution and what they elaborate turns out to be dual. Each of them, in fact, chooses to reject a rule of classical logic involved in the argument: on the one hand, Field rejects the rule of the introduction of the conditional; and, on the other hand, Priest rejects the rule of the elimination of the conditional, i.e. *modus ponens*. Moreover, both logicians don't completely reject the two rules, but, instead, they restrict them to classical contexts. In the case of Field's theory, these contexts correspond to those in which excluded middle holds, whereas in the case of Priest, they correspond to consistent contexts, where the *ex contradiction quodlibet* holds.

4.3.4 Acceptance and Rejection

Another aspect of the comparison between the paracomplete theory of Field and the dialethic theory of Beall concerns the notions of acceptance and rejection. Acceptance and rejection are traditionally taken as cognitive states. In particular, accepting something is nothing but believing that thing, and rejecting something is refusing to believe that thing. There are two speech acts corresponding to those cognitive states, that is, certain linguistic acts we use to say when we have the corresponding cognitive states. The speech acts for acceptance and rejection respectively are assertion and denial. They are strictly connected to their cognitive states: asserting something means that one accept that thing, and denying something means that one rejects that thing.

In order to better explain the behaviour of acceptance and rejection, Field proposes to consider the notion of "degree of belief" that we can use to translate the two cognitive states. In this sense, accepting something corresponds to having a degree of belief over some threshold τ ,⁴³ and rejecting something corresponds to having a degree of belief lower than the corresponding co-threshold, $1 - \tau$.⁴⁴

According to classical logic, acceptance and rejection are interdefinable in a way that the rejection of a claim is equivalent to the acceptance of the negation of that claim. In this view, rejection is a derivative notion. This, in terms of degree of belief means that the sum of the degrees of belief in a sentence and in its negation is equivalent to 1. Therefore, in classical logic we have that all sentences are cases of what we can call *full acceptance* or *full rejection*, that is, the threshold is always equal to 1 and the co-threshold is always equal to 0. This, in fact, can be shown in a very simple way, through the introduction of a general law for degrees of belief. This law involves a specific notion of subjective probability and, for this reason, is called *the general law for classical subjective probability theory*. Its formulation is the following:

$$(SP) \quad P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$

⁴³ "Over" in this definition can be understood in two different ways: "greater than" or "greater than or equal to", depending on the context.

⁴⁴ Field[2008], p.74.

where P is the agent's degree of belief – or, equivalently, the subjective probability – function. In the specific case where B is $\neg A$, SP is equivalent to what follows:

$$P(A \vee \neg A) = P(A) + P(\neg A) - P(A \wedge \neg A)$$

that, in turn, is equivalent to the following:

$$P(A) + P(\neg A) = P(A \vee \neg A) + P(A \wedge \neg A)$$

Since we know that the excluded middle is accepted by all classical logicians, we have that the degree of belief in its instances is 1, that is, we know that $P(A \vee \neg A) = 1$. We also know that in classical logic contradictions are not admitted, or, in other words, we know that the classical logicians reject the existence of sentences such that they and their negation are both true. This means that in classical logic we have that $P(A \wedge \neg A) = 0$. If we substitute these values into the above equation, we obtain $P(A) + P(\neg A) = 1$, that means nothing but that, according to classical logicians, the sum of the degrees of belief of an agent in a sentence and in its negation is equal to 1.

However, things are different for non-classical logicians. In fact, both paracompleatists and dialetheists deny that acceptance and rejection are interdefinable, that is, they deny that the rejection of a sentence is equivalent to the acceptance of the negation of that sentence.⁴⁵ There is a thing classical and non-classical logicians agree on, namely that acceptance and rejection are exclusive but non-exhaustive notions. As we already saw in section 1.3.2, classical logicians maintain that the two notions are incompatible and, for this reason, they are exclusive. Furthermore, they think that the acceptance and rejection are not exhaustive notions because it is always possible to take an agnostic – or, neutral – position about a sentence.

With regard to the paracomplete theory developed by Field, as we said, it's important to make a distinction between rejection and acceptance of the negation. In fact, according to Field, they are not equivalent to one another. In particular, as Field puts it:

Rather than explaining rejection in terms of acceptance (as in *commitment not to accept*) we should regard acceptance and rejection as dual notion.
[Field 2008, p.74]

Therefore, according to Field, we have that rejecting a sentence A doesn't automatically lead to the acceptance of its negation $\neg A$, because, on the one hand, one can always reject both A and $\neg A$, but, on the other hand, it still remains impossible to accept and reject the same sentence at the same time.

If we take into account the notion of degree of belief, we obtain that in Field's theory the sum of the degrees of belief of an agent in a sentence A and in its negation $\neg A$ don't have to be necessary equal to 1, contrary to classical logic, but can also be lower. If we consider, in fact, the general law for classical

⁴⁵ Beall[2009], p.51 and pp.100-101, and Priest[1987], p.98.

subjective probability theory, we can show very well this result. As a matter of fact, we know that Field's paracomplete logic doesn't admit the existence of dialetheias, hence, we have that $P(A \wedge \neg A) = 0$, that is, we have that the degree of belief in a contradiction is always equal to 0. On the other hand, we also know that excluded middle is not universally valid in Field's logic. This means that the degree of belief of an agent in an instance of the excluded middle can be lower than 1, or, using the symbolism, we have $P(A \vee \neg A) \leq 1$. Hence, in a paracomplete logic, we have the following result:

$$P(A) + P(\neg A) \leq 1$$

In other words, the sum of the degrees of belief of an agent in a sentence and in its negation can be lower than 1. Hence, in a paracomplete logic, we can have rejection without having acceptance of the negation as well. Speaking in terms of degrees of belief, in a paracomplete theory we can have $P(A) \leq 1 - \tau$ without having $P(\neg A) \geq \tau$.⁴⁶

We have already talked about Priestian account of rejection and acceptance in section 1.3.2. It's important to underline that Beall shares exactly the same non-classical account of acceptance and rejection endorsed by Priest.⁴⁷ By using Beall's words:

While nobody accepts and rejects something at the same time, acceptance and rejection are properly modeled non-classically. [Beall 2009, p.100]

Let's recall, now, the main features and develop the discussion through the aid of the new concepts just introduced.

As we saw, Priest (and so Beall), likewise Field, deny that acceptance and rejection are interdefinable. However, the reasons are different from those abducted by the paracomplete philosopher. While according to Field we can reject both a sentence and its negation, according to Priest and Beall we can accept both a sentence and its negation, that is, we can accept both A and $\neg A$. Despite this difference, they agree on an important point: according to both philosophers, it is always impossible to accept and reject the same sentence at the same time, because the two notions are incompatible to each other. This

⁴⁶ Talking about degrees of belief is useful to distinguish Field's paracomplete theory from classical gap theory. As we already noticed, according to classical logic and, in particular, according to classical gap theory, the sum of an agent's degrees of belief in a sentence and its negation is always equal to 1, that is, in classical theory we have $P(A) + P(\neg A) = 1$. In fact, classical gap theorists believe that $P(\text{True}(A)) + P(\text{True}(\neg A)) \leq 1$, that is, they believe that the sum of degrees of belief in the attribution of truth to a sentence and in the attribution of truth to its negation can be lower than 1. However, the point here is that the degree of belief in $\text{True}(A)$ – and $\text{True}(\neg A)$ as well – can be lower than the degree of belief in A – and $\neg A$ as well –, or, in other words, $P(\text{True}(A)) \leq P(A)$ and $P(\text{True}(\neg A)) \leq P(\neg A)$. In Field's paracomplete theory there isn't this difference at all. As a matter of fact, according to Field we have that $P(\text{True}(A)) = P(A)$ because in Field's theory the full T-schema holds, whereas in classical gap theory it doesn't. In the latter theory, in fact, only one side of the T-schema – namely, the left to right side, i.e. $T(\langle a \rangle) \rightarrow a$, but not the right to left one, i.e. $a \rightarrow T(\langle a \rangle)$ – holds.

⁴⁷ Beall[2009], p.51 and pp.100-101.

means that, according to the dialetheist, we can accept $\neg A$ without being forced to reject A .

Thanks to the notion of degree of belief we can clarify even more the dialethic idea about this issue. According to dialetheists, similarly to what Field claims, the sum of the degrees of belief in a sentence and in its negation is not necessary equal to 1. While in Field's theory the sum can be lower than 1, in Priest and Beall's paraconsistent theory can be higher. If we take into account the general law for classical subjective probability theory and we apply it to dialetheism, we obtain a result that mirrors the one obtained by the paracomplete theory. First, we know that in paraconsistent logic the law of excluded middle always holds, hence, we have that an agent's degree of belief in an instance of the excluded middle must be 1, that is, $P(A \vee \neg A) = 1$. Moreover, we know that dialetheism admits by definition the existence of true contradictions, hence, the degree of belief in a contradiction can be higher than 0 in Priest and Beall's theories, that is, $P(A \wedge \neg A) \geq 0$. Therefore, the result for a paraconsistent logic is the following:

$$P(A) + P(\neg A) \geq 1$$

This means that the sum of an agent's degrees of belief in a sentence and its negation can be higher than 1, according to dialetheism. Hence, in a paraconsistent logic, such as the one assumed by dialetheism, we can accept the negation of a sentence without being forced to reject the sentence itself. In other words, we can have $P(\neg A) \geq \tau$ without being forced to have also $P(A) \leq 1 - \tau$.⁴⁸

What the analysis has highlighted, thus, is that both Priest – as well as Beall – and Field have a non-classical account of rejection, because they maintain that rejection is a primitive notion that is not interdefinable with acceptance. Moreover, the views of the two philosophers about the notions of acceptance and rejection turned out to mirror each other: on the one hand, according to Field, rejecting a claim doesn't force us to accept its negation; on the other hand, according to Priest, accepting the negation of a claim doesn't force us to reject that claim. If we also use the formalism related to the concept of degrees of belief, the duality between the two accounts appears even more clear: according to Field, the sum of an agent's degrees of belief in a sentence and in its negation can be lower than 1, i.e. $P(A) + P(\neg A) \leq 1$; contrariwise,

⁴⁸ Similarly to what we said about classical gap theory, we can now make a difference between dialetheism and classical glut theory through the use of the notion of degree of belief. As we already noted, according to classical logic and hence, in particular, according to classical glut theory, the sum of an agent's degrees of belief in a sentence and in its negation is always equal to 1, that is, $P(A) + P(\neg A) = 1$. Classical glut theorists clearly believe that $P(\text{True}(A)) + P(\text{True}(\neg A)) \geq 1$, that is, they believe that the sum of degrees of belief in the attribution of truth to a sentence and in the attribution of truth in the negation of that sentence can be higher than 1. However, the point here is that the degree of belief in $\text{True}(A)$ – and $\text{True}(\neg A)$ as well – can be higher than the one in A – and $\neg A$ as well –, or, in other words, $P(\text{True}(A)) \geq P(A)$ and $P(\text{True}(\neg A)) \geq P(\neg A)$. This gap between the degrees of belief in a sentence and in the attribution of truth to that sentence is a peculiar feature of classical glut theory. According to dialetheism, things are different. According to dialetheists, $P(\text{True}(A)) = P(A)$ because they validate the full T-schema, contrariwise to classical glut theorists, which accept the right to left side of the T-schema and reject the other side, i.e. $T(a) \rightarrow a$.

according to Priest, we have the opposite situation, namely that the sum of an agent's degrees of belief in a sentence and in its negation can be higher than one, i.e. $P(A) + P(\neg A) \geq 1$. Finally, what is worth noting is that the special situations where the value is lower than 1 in a paracomplete logic are exactly the same cases where the value is higher than 1 in a paraconsistent logic, and, more specifically, they correspond to defective sentences, such as the liar and its kin.

With regard to this topic, the conclusion is that there seems to be no *ad hoc* reason why to choose dialetheism over Field's paracomplete theory. In fact, the two theories turn out to be equally good with regard to the pragmatic account of the notions of acceptance and rejections they provide, because they propose two very similar accounts under several aspects and, as far as the differences are concerned, they provide two totally specular views.

4.3.5 The Revenge Phenomenon

In section 1.5.1 we defined the so-called revenge phenomenon in the following way: the revenge phenomenon is a particular reformulation of the semantic paradoxes, such as the liar, that uses the key notions of the solutions to those paradoxes provided by different theories. Or, as Priest puts it:

There is a certain notion the intelligibility of which the theorist presupposes which, if it is included in the language in question, can be used to refashion the paradox. [Priest 2005, p.44]

In other words, this means that wherever a theory provides a solution to the liar, the notions used to formulate the solution are used to construct a new and strengthened liar. As a result, we will have as many revenge liars as the theories providing a solution to the standard liar. Since every semantic theory faces the problem of the liar and since a semantic theory is adequate only if provides a solution to that paradox, then, as a consequence every adequate semantic theory of truth will face the problem of the revenge phenomenon of the liar. As Priest explains the issue:

“Extended paradoxes” are simply the results of applying the construction in different theoretical frameworks. [Priest 2007, p.226]

The target of this section is trying to figure out whether dialetheism and Field's paracomplete theory are subject to the revenge of the liar, or not, and, if this is the case, what are the solutions they provide. My purpose here is to see whether, as long as the revenge phenomenon is concerned, it's possible to establish which one between dialetheism and the paracomplete theory is the preferable, or not, that is, whether the solution to the revenge liar provided by the former is better than the one of the latter, and vice versa, or not. In other words, I'm going to try to answer the following question: does the revenge liar make a difference in the analysis we are providing?

Before we analyse non-classical theories of truth, we should very briefly take into account classical ones, in order to see what the solution they put forth to the revenge of the liar is. Traditionally, classical theorists developed a solution for the revenge liar that involves a restriction in the language of the

theories. In fact, they deny that the new concepts – namely, those involved in the solution of the standard liar and, hence, those involved in the formulation of the revenge liar – are expressible in the language of the theory.⁴⁹ If it's true that this would solve the problem, it is also true that the price we must pay for this move would be very high. As a matter of fact, saying that the notions involved in the formulation of the revenge liar are not expressible in the language of the theory means also saying that those paradox-solvers are talking in a language different from the language for which the semantic is offered, and this is not a desirable result for a theory of truth.⁵⁰

Let's now come back to non-classical theories of truth and let's start the comparison by calling to mind the solution to the revenge liar put forth by Priest.⁵¹ According to Priest, the revenge phenomenon represents a problem only for those theories, classical or not, that mean to preserve consistency to all costs. The reason is that consistency can be maintained only at the cost of incompleteness, that is, only denying that the key notions involved in the formulation of the revenge liar are expressible in the language of the theory. Since dialetheism doesn't have consistency within its targets and, on the contrary, inconsistency is the very thesis dialetheists affirm, then, according to Priest, dialetheism is invulnerable to the revenge of the liar.

The first revenge liar that naturally arises for dialetheism is the following:

(R) (R) is false only.

The evaluation of (R) leads dialetheists to the conclusion that it is both true and false only. However, this seems to be an insoluble contradiction also for the dialetheist. Of course, Priest and Beall don't agree with this diagnosis. According to them, (R) has an immediate and very intuitive solution that is perfectly in line with the spirit of dialetheism, namely, (R) is both true and false, and false only. Hence, the adverb "only" – and consequently the truth-value "false only" – only appears to have an exclusive value, because, according to Priest and Beall, it doesn't, since it is perfectly compatible with whatever other designated one.

It's important to notice that, according to Beall, the value "true only" is perfectly equivalent to the value "ttrue", that remember is the name for the transparent truth predicate in Beall's theory. The reason is that in a deflationary theory of truth there is no difference between the two values. Let's see how Beall's argument proceeds.

⁴⁹ I'm going to take the problem of the expressive power of the theories into account in next section. For now, I just want to say that the two problems – i.e. the revenge phenomenon and the expressive limitedness of the language – are strictly connected to each other, so that the former seems to have the latter as an inevitable consequence, and vice versa.

⁵⁰ This problem in first place affects, as we already know, the theory developed by Tarski, which postulates the existence of a hierarchy of languages, where every language in the hierarchy can contain only the truth predicate for the previous language and it cannot contain its own truth predicate. Moreover, as it's well known, despite all his efforts, also Kripke's theory is affected by the same problem, because inside the theory we cannot in anyway say that the liar is not true, since there is no fixed point containing the liar. As Kripke himself recognizes:

The ghost of the Tarski hierarchy is still with us. [Kripke 1975, p.714]

⁵¹ See section 1.5.1 for the details on the revenge liar for dialetheism.

The value “true only” applied to a sentence says that that sentence is true and it is not the case that its negation is true as well. It can be formalized in the following way:

$$T(\langle\alpha\rangle) \wedge \neg T(\langle\neg\alpha\rangle)$$

but, remember, that in a deflationary theory of truth the truth predicate is transparent. This means that the above formalization of “true only” is equivalent to what follows:

$$\alpha \wedge \neg\neg\alpha$$

by the principle of double negation, this is equivalent to α , and hence to $T(\langle\alpha\rangle)$ because of the transparency of truth. Therefore, “true only” is equivalent to “ttrue”. In particular, Beall wishes to underline that:

The thought [...] that we need to recognize a notion of truth stronger than ttruth seems to me to be in need of argument. [Beall 2009, p.58]

This means that “true only” doesn’t bring any new problem that “ttrue” doesn’t already have. If “true only” is equivalent to “ttrue”, then (R) is not a genuine revenge liar, because it is equivalent to the standard liar and it can be solved in the same way. After all, as Beall says:

We did not introduce the device [i.e. ttrue] so as to name some important property; we brought it in only for practical, expressive convenience. [...] When, in the face of various spandrels, we recognize gluts, we are recognizing more ttruths; we are not suddenly bringing in a predicate “glut” to name an important, non-logical property. The “category” of *gluts* is itself derivatively expressive; it is shorthand for ttruths that have true negations. Why expect, then that the notion of “non-glut” or “true only” (i.e. true non-glut) is any different? As above, I think it’s not different: *just true* just is *ttruth*. [Beall 2009, p.51 (emphasis in original)]

As we will see, this solution comes with exactly the same problems that a solution provided by non-deflationary dialetheists brings with it. Let’s now say little more on non-deflationary dialetheic accounts, before we can look upon these problems.

As we said with regard to the previous revenge liar, the solution provided by Priest and Beall seems not to be totally satisfactorily because it seems to violate the meaning of the adverb “only”. In fact, in a dialetheic view, the value “true only” is perfectly compatible with any other designated value. There is a bunch of attempts to construct a genuine revenge liar that also a dialetheist can’t solve assuming that the notions involved in its construction has their usual meaning. For instance, in section 1.5.1, we saw the attempt by Simmons and Littman that propose to consider the sentence (Z) that says of itself that it is v , where v is the value correspondent to an arithmetic falsity, such as “1+1+3”. Since arithmetic is almost universally supposed to be consistent, then (Z) seems to be a genuine revenge liar for dialetheism. However, as we saw,

Priest is not of the same idea. In fact, he considers (Z) as both true and v . Therefore, there seems to be no exclusive truth-value for a dialetheist or, in other words, there seems to be no truth-value that is incompatible with a designated one.

The main problem with this approach is that it seems that we can always add a designated truth-value to a non-designated value, even if the latter seems to prevent it because of the meaning of the words involved. This, by itself, may not be a big deal, but it becomes a problem because it directly leads to the related problem we will see in next section, namely the problem of the expressive limitedness of the language of the theory. In fact, if “false only” and its kin are meant to have always a non-exclusive value, how is it possible to express disagreement with a claim by using the dialethic logic, provided that disagreement actually requires that exclusivity?⁵² Therefore, it seems that the revenge liar is a real problem for a dialethic theory because either the dialetheist admits that her theory is not immune to the revenge of the liar –and, by so doing, she admits that her theory is not appropriate to do the job which it was developed for, i.e. to solve semantic paradoxes – or she solves the liar by denying that there is an exclusive truth-value – or, in other words, by denying that there is a semantic concept that is genuinely incompatible with truth – and, by so doing, slipping back into the problem of the incompleteness of the language.

Now, what about Field? The first aspect shared by the two philosophers is that Field, as well as Priest, claims his theory to be immune from the revenge of the liar:

The new conditional can be used to show that the theory is not subject to “revenge problems”. [Field 2003a, p.140]

The solution to semantic paradoxes provided by Field is called by Field himself *G-solution*.⁵³ There is not a real meaning associated with this name, but Field reassures us by saying:

If you want to think of the “G” as standing for “good” I won’t stop you.
[Field 2007, p.78]

G-solution is particularly appreciable, according to Field, because is able to keep the full T-schema. In order to achieve this requirement, Field is forced to sacrifice the law of excluded middle, as we know, by restricting it to all non-paradoxical contexts. In keeping with these desiderata of Field’s theory there is the one holding the rejection of the material conditional, that is, the rejection of the equivalence between $A \rightarrow B$ and $\neg A \vee B$. The reason is that by rejecting the material conditional we can accept the biconditional $T(\langle Q \rangle) \leftrightarrow \neg T(\langle Q \rangle)$ – where (Q) is the liar sentence – that directly follows from the T-schema. If the conditional were the material one, we would have rejected the previous biconditional, since it would have been an instance of the excluded middle, and

⁵² Also in Beall’s account we face the same problem. In fact, if “true only” is perfectly equivalent to “true”, how can we explain the difference between the two values? And, in particular, how can we preserve the exclusive character of the former? There seems no way for “true only” to express this exclusivity.

⁵³ Field[2007], p.79.

excluded middle doesn't hold in Field's theory. Moreover, as we saw in section 4.3.1, G-solution doesn't admit the existence of dialetheias.

The problem of the revenge of the liar seems to arise from the possibility in Field's theory to produce a kind of never-ending hierarchy of liar sentences, each of which seems to be more paradoxical than the previous. In particular, Field can explain the defective status of the liar only via the introduction of a determinacy operator that allows him to say that the liar is neither determinately true nor determinately false. The problem becomes now clear: this solution inevitably leads to the arising of new paradoxes, that is the revenge phenomenon of the liar. In particular, there will be a sentence that says of itself that it is not determinately true. How can we consistently deal with that sentence? What is certain is that an adequate solution to the semantic paradoxes doesn't have to take into account only sentences involving the truth predicate alone, but also those sentences involving the determinacy operator as well.

Field claims he can consistently and satisfactorily treat every level of the hierarchy exactly by using his G-solution. As he puts it:

The existence of these hierarchies prevents certain kinds of revenge problems from arising: certain attempts to state revenge problems simply involve going up a level in a hierarchy of which have been given a non-paradoxical treatment. [Field 2007, p.79]

If the hierarchy allows us to solve each stronger revenge liar that arises, it seems that the only way we have to make a genuine revenge liar for Field's paracomplete theory – that is, a revenge liar that the hierarchy cannot deal with – is by breaking out of the hierarchy. Is this operation possible? On the one hand, it would be better if it wasn't, so that we can avoid the revenge of the liar; on the other hand, however, if we cannot break out of the hierarchy, then we will face the same problem that all the other theories face, dialetheism included, namely, an expressive limitation of the language of the theory.

More in details, G-solution can solve the problem of the liar and its revenge through the introduction and, then, the iteration of the determinacy operator. As Field puts it:

Thus with “determinately” defined as D, it is immediate that there can be no new paradoxes of determinacy. [Field 2007, p.111]

In other words, every single liar of the hierarchy is solved at the following level thanks to the iteration of the determinacy operator. This is possible because excluded middle doesn't hold for determinate truth, that is, it's not the case that every sentence is either determinately true or not determinately true.

Let's now consider the liar hierarchy, whose behaviour can be summarized as follows:

$(Q_0) \quad \neg T(\langle Q_0 \rangle)$
 $(Q_1) \quad \neg DT(\langle Q_1 \rangle)$
 $(Q_2) \quad \neg DDT(\langle Q_2 \rangle)$
 and so on.

Field's solution is to individually solve each liar of the hierarchy by applying the determinacy operator one more time at each stage. In this way, (Q_0) will have the following solution: $\neg DT(\langle Q_0 \rangle)$ and $\neg D\neg T(\langle Q_0 \rangle)$. However, this provides us the means for a new liar, that is, a sentence that says of itself that it is not determinately true, and this sentence is exactly (Q_1) . To solve (Q_1) , we must iterate one more time the determinacy operator such that the solution turns out to be $\neg DDT(\langle Q_1 \rangle)$ and $\neg DD\neg T(\langle Q_1 \rangle)$. The solution provided for (Q_1) , in turn, allows us to construct a new liar, namely, a sentence that says of itself that it is not determinately determinately true, i.e. (Q_2) , and so on. The result, in this way, is a hierarchy of liars that can be solved only at the following level. To be more technical, Field takes into account a transfinite liar hierarchy, such that every liar sentence, (Q_σ) , is equivalent to $\neg D^\sigma T(\langle Q_\sigma \rangle)$ and where D^σ is obtained by transfinitely iterating the determinacy operator D . The solution proposed by Field is to add one more time the determinacy operator in order to solve each liar sentence of the hierarchy, (Q_σ) , by saying $\neg D^{\sigma+1} T(\langle Q_\sigma \rangle)$ and $\neg D^{\sigma+1} \neg T(\langle Q_\sigma \rangle)$.⁵⁴

For all this reasons, it's possible for Field to solve the problem of the revenge liar only ascending through the levels in the hierarchy of determinacy operators, so that he can say that a liar sentence is not determinately true at the following level. We can extend the iteration process into the transfinite and this, according to Field, does not generate new paradoxes. In fact, the hierarchy never collapses and this means that the hierarchy "becomes strictly stronger for as long as the iteration is satisfactorily definable."⁵⁵ Therefore, to say that the hierarchy never collapses doesn't mean that it goes on forever. In fact, this cannot be because the language contains only countable expressions. There must be, thus, a countable α for which there isn't a correspondent determinacy operator in the language, that is, there isn't D^α . In conclusion, the hierarchy never collapses means that it continues for a long way through the countable ordinals.⁵⁶

If it is not possible for the liar to vindicate itself in this way, is there another possibility? Priest claims there is.⁵⁷ It seems, in fact, possible to introduce a new operator, Δ , allowing the expression of the *general* notion of determinate truth and this is where the liar revenge is at play. As Priest puts it:

None of the D^σ predicates expresses determinate truth in general; and it is this that gives rise to the paradigm revenge problem [Priest 2005, p.45]

Even if no determinacy operator within the hierarchy allows the expression of determinate truth in general – because we can only say of each individual liar of the hierarchy that it is not determinately true –, it is always possible to add an operator to the language that performs this function. Such operator, Δ ,

⁵⁴ This is a rough simplification in order to give the intuitive idea of how the iteration of the determinacy operator in Field's theory works. Field gives the conditions the determinacy operator must satisfy. Moreover, he shows how the determinacy operator works and how this allows to solve each level of the transfinite liar hierarchy through the aid of some explicative examples. For the technical details see Field[2007], pp.29-33 and Field[2008], section 15.2.

⁵⁵ Field[2008], p.115.

⁵⁶ Ibid., pp.237-238.

⁵⁷ Priest[2010], p.124.

guarantees us that A is indeterminate when we have the following: $\neg(\Delta A \vee \Delta \neg A)$, that is equivalent to $\neg\Delta A \wedge \neg\Delta \neg A$. Hence, we can now construct a revenge liar for Field's paracomplete theory:

$$(F) \quad \neg\Delta F$$

it's easy to see that from (F) a contradiction follows:

- | | | |
|-----|--|--|
| (1) | $\vdash \Delta F \rightarrow F$ | [By definition of ΔF] ⁵⁸ |
| (2) | $\vdash \Delta F \rightarrow \neg\Delta F$ | [by 1 and equivalence between F and $\neg\Delta F$] |
| (3) | $\neg\Delta F$ | [by 1,2 and LEM] |
| (4) | F | [by the equivalence between F and $\neg\Delta F$] |
| (5) | ΔF | [by 1 and 4] |
| (6) | $\Delta F \wedge \neg\Delta F$ | [by 3, 5 and \wedge -Introduction] |

Therefore, a contradiction immediately arises from (F). The obvious move we expect from Field is to deny the excluded middle to hold for (F). Excluded middle doesn't hold for a sentence only if this sentence is indeterminate. Hence, in the specific case, (F) is not determinate. However, as we saw, to say that (F) is indeterminate means nothing but what follows: $\neg\Delta F \wedge \neg\Delta \neg F$. The problem is that this is equivalent to an instance of excluded middle that we have just denied to hold. As a matter of fact, from $\neg\Delta F \wedge \neg\Delta \neg F$ follows $\neg\Delta F$ by \wedge -Elimination, and from $\neg\Delta F$ follows $\Delta F \vee \neg\Delta F$ by \vee -Introduction. The only way to avoid this contradictory conclusion is to admit this general notion of determinateness, expressed by the operator Δ , not to be in the language.

To sum up this section, both dialetheists and paracompletists maintain their own theories to be immune to the problem of the revenge of the liar. This turned out to be true for both the theories only provided they abandon a very important requirement for a theory of truth, namely the expressive completeness of the language. Thus, it seems that in both cases – and we will see it in more details in next section – it's not possible to have both an adequate solution to the revenge problem and the expressive completeness of the language of the theory. From this point of view, hence, there is no advantage of a theory over the other. In fact, Field himself, as far as the revenge problem is concerned, maintains:

However, I also think that any *apparent* revenge problem for a paracomplete theory of Part III arises *with at least equal force* for paraconsistent dialethic theories: there is no possible advantage of paraconsistent dialethic theories over paracomplete theories with regard to revenge. [Field 2008, p.384]

⁵⁸ (1) holds because $T(\langle DF \rangle)$ is an instance of ΔF and we have $T(\langle DF \rangle) \rightarrow F$ by the intersubstitutivity principle.

4.3.6 The Expressibility Problem

As it came up in the previous section, both dialetheism and Field's paracomplete theory face the same dilemma: either they provide a solution to the problem of the revenge of the liar, or they keep the expressive completeness of the language of their theories. One or the other, it seems not possible to have both. To restate the dilemma by using Priest's words, what we want for a theory of truth is the satisfaction of the following desiderata:

[a] we must be able to say of certain sentences, e.g. the liar, that they are of this kind [defective]; and [b] we must be able to talk about such sentences in general and say things about them. [Priest 2010, p.122]

We can have either [a] or [b], but not both.

As we saw, the problem for Field's theory is to express the general notion of determinate truth. Field is able to express within his theory the defective status of each liar that arises. In fact, it's sufficient to move up to the next level through the iteration of the determinacy operator. However, what Field is not able to do within his theory is to say of a sentence that it is not determinately true in general, because, as we saw, if he was able to express that notion, this would give rise to a strengthened liar that cannot be solved. Hence, the only possibility for Field is to deny this general notion of determinate truth to be expressible in the language of the theory and, in so doing, he is forced to admit what follows:

My claim is that the notion is ultimately unintelligible. [Field 2008, p.356]

Hence, the result is that in Field's theory is not possible to express the key notions of the theory itself. In other words, one of Field's targets was to express the defective status of certain problematic sentences, such as the liar. This is possible via the definition within the language of a determinacy operator that allows the claiming that the liar is neither determinate true nor determinate not true. To this extent, we are able to satisfy the requirement [a] from Priest's quote, because the determinacy operator enables us to say that the liar belongs to the defective sentences. However, the result is the abandon of Priest's second requirement that as a consequence has an important expressive limitation for the language of the theory. As a matter of fact, as long as KFS is concerned, Field himself says what follows:

The inability of KFS to express a notion of determinateness is a crippling limitation. [Field 2008, p.78]

This, of course, applies to every theory of truth, also to the one developed by Field that, for this reason, faces major expressive problems. As Priest emphatically puts it:

To declare all general claims about indeterminacy unintelligible is an act of ladder-kicking-away desperation of *Tractarian* proportions. [Priest 2010, p.124]

The only possibility for Field to avoid the revenge paradox is, then, to declare the things he cannot express in the language of his theory to be meaningless, and this is a high price to pay. In fact, the notions we have to ban from the language – such as determinacy – appear to be perfectly intelligible but they cannot be expressed in the theory, on pain of contradiction.

Are things different for dialetheism? In a sense yes, they are different and in another sense they are not. Let's clarify the issue. As we saw, dialetheism faces a bunch of problems strictly tied to the expressibility of certain key notions of the theory. In section 1.5.2 we saw that the dialetheist is accused of being able to solve the strengthened liar only at the cost of abandoning the possibility of expressing in the language of the theory its distinctive notions, that is, in particular the notions of "true only" and "false only". Priest actually claims he can express that a sentence is only false – that is, false and not a dialetheia – and he can do that exactly by using these very words. What the dialetheist cannot do is to assure that what he is saying has a consistent behaviour.⁵⁹ However, as Field puts it, this is not what we meant to achieve:

The problem that these authors are raising is that on these definitions, the notions don't behave in accordance with how they seemed intended to behave when the theory was being explained. [Field 2008, p.386]

In other words, Priest can express in the language of his own theory the truth-value "only false" – in the sense that he has the words inside his language to do this job – but he cannot in any way ensure that this value is exclusive. But this is exactly the guarantee we were looking for when we used the adverb "only".

Moreover, always assuming that Priest is right on this point, there is a further issue to be addressed by the dialetheist, that is strictly linked to this expressive problem and is not so easy to solve. How can dialetheists express disagreement? In other words, how can dialetheists express the exclusivity the value "only false" should have? If I say "A", how can dialetheists express disagreement with my assertion? They cannot, in fact, say " $\neg A$ " to express disagreement, because in a paraconsistent logic it could be the case of both A and $\neg A$, and, hence, this wouldn't be a genuine disagreement. Neither they can express disagreement by saying "A is false" or "A is false only", because, as we just saw, neither is exclusive. From a logico-semantic point of view, dialetheists have no way to express disagreement – or, exclusivity – inside the language of their theory. In other words, from a logico-semantic point of view, dialetheists are in the same position of paracompletists, that is, they are not able to express the key notions of their own theory. This is why I claimed that things are not different for dialetheism compared with Field's paracomplete theory.

Priest claims the dialetheist has an advantage over the paracompletist. While the latter can in no way express the notions of defectiveness and determinateness – that are the distinctive notions of her theory –, the former has a way to express her own key notions – that is the notion of dialetheia – even if she cannot do it in a consistent way. I don't think this is a genuine advantage of

⁵⁹ Priest maintains that nobody – neither a dialetheist nor a classical logician – is able to provide this guarantee. Even if a classical logician says "A is false" we cannot be sure that A is not true as well. The difference is that if that was the case, classical logic will explode into trivialism, whereas this danger is averted by dialetheism via the assumption of a paraconsistent logic.

dialetheists over paracompletists, because, as we said, there is another notion that directly arises from Priest's theory and cannot be expressed in it by keeping its traditional exclusive meaning safe, namely the notion of "false only". In fact, Priest can express the notion of "true false", only provided that a change in the meaning of the words must be done, that, in other words, means that he gives to the words a different meaning from its usual one.

However, there is a way available to the dialetheist for expressing disagreement, but involves notions belonging to the reign of pragmatics and, for this reason, it doesn't affect the analysis I'm conducting here.⁶⁰ I'm referring to the notion of "denial" and its correspondent cognitive state "rejection". Rejection and acceptance, as we stressed in different sections, are not interdefinable notions because, according to dialetheists, we can accept both a sentence and its negation at the same time. Moreover, what allows the dialetheist to solve the problem is the exclusive character of the two notions: acceptance and rejection are incompatible notions, that is, we cannot both accept and reject the same sentence.

Also Beall counts on the reign of pragmatics in order to solve the problem of disagreement. The difference between Priest and Beall is that the latter doesn't use a pragmatic operator, but, instead, his account is based on conversational implicatures. As Beall himself puts it:

Perhaps "just true" carries pragmatic implicatures not carried by (an utterance of) "ttrue", something to the effect that the speaker rejects the given negation. [Beall 2009, pp. 51-52]

By so doing, we can achieve our purpose. On the one hand, "only true" – or "just true", it's the same – and "ttrue" are still equivalent, like they turn out to be in a dialethic and deflationary theory of truth.⁶¹ On the other hand, we are also able to express the intuitive difference between the two truth-values, that is, we are in some way able to express the exclusive character of "only true" that "ttrue" doesn't have: the former carries certain pragmatic implicatures don't carried by the latter. From a semantic point of view the two values are equivalent, but they convey different pragmatic conversational implicatures that allows "only true" to perform its role of expressing disagreement value.

This seems to solve all the problems faced by the dialetheist. On the one hand, it is not possible to construct a liar sentence by using "denial", because the latter is a pragmatic and not a logical operator and, for this reason, is a force-operator that applies to the entire sentence and doesn't interact with the content of the utterance. On the other hand, the incompatibility between acceptance and rejection allows the dialetheist to express disagreement, namely by denying (rejecting) the sentence she disagrees with.

However, it is precisely because "denial" is a force-operator that it faces a bunch of problems that disqualify it for the role of expressive device for disagreement. We can avoid the revenge of the liar thanks to the fact that "denial" is a force-operator, but this is exactly what limits the general applicability of the device. In fact, "denial" cannot be used in richer embedded

⁶⁰ Because, remember, my purpose is to focus only on the logico-semantic level.

⁶¹ See section 4.3.5.

sentences, because it doesn't have its intended meaning since it applies only to the whole sentence.

To sum up, it's true that dialetheism has a way to express disagreement, but, on the one hand, this is a very limited solution from its applicability point of view and, on the other hand, it lies in the reign of pragmatics and not in the logico-semantic one. Moreover, Priest himself maintains that this pragmatic way is viable also by Field because also Field can express that a sentence is not determinately true using "denial", in the same way as Priest does. Therefore, also from this point of view seems not to be any advantage of a theory over the other. As Priest puts it:

Field has no way of asserting that something is not true, if this is meant to include things that are false and things that are indeterminate; I cannot assert that something is false-only if this is required to exclude things that are true as well. For both of us, though, there is a way of obtaining this effect with a different kind of speech act, namely denial: both of us can deny that A is true. [Priest 2010, p.136]

As far as dialetheism is concerned, denying a sentence A means expressing a genuine disagreement with A, because it assures us that one is not accepting A as well. According to Field, on the contrary, denying A is not incompatible with the denial of $\neg A$ and, hence, denying A means that A can be either false or indeterminate and with a denial of A with don't know which category A belongs to.

In conclusion, thus, my target in this section was to see whether there is a logico-semantic way to express the key notions of the two theories, or not. The result is that this is not the case neither the dialetheist nor the paracompletest can provide a decisive benefit of her theory over the rival's one. Actually, the analysis showed that the two theories mirror each other also if we move from logic to pragmatic, because the solution proposed by Priest if works for dialetheism, then it will work also for Field's theory. In any case, we saw it isn't a satisfactory solution even from a pragmatic point of view. Hence, both theories are facing a dilemma: either they solve the problem of the revenge of the liar by severely limiting the expressive power of the language of their theories, or they keep the expressive completeness of the language by giving up the idea of finding out a solution to the revenge liar. Either case, neither of them turns out to have a decisive advantage (or disadvantage) over the other as far as the logico-semantic structure of the theories is concerned.

4.3.7 Conclusion

The analysis I have just concluded doesn't enable to establish which one between dialetheism and Field's paracomplete theory is the best. Indeed, the aspects I have considered allows to claim that Beall's dialetheism and Field's theory mirror each other, that is, it turned out they are perfectly dual, at least from a logico-semantic point of view. Both are non-classical theories of truth that validate the intersubstitutivity principle of the truth predicate. Both are three-valued logics and the matrix of their truth tables is the same, what changes is the interpretation they give to the third truth-value that in dialetheism is "both true and false" and in Field's theory is "neither true nor

false”. Moreover, it’s possible to deduce both theories from the same starting point, that is, we can deduce them depending on the interpretation of Kripke’s construction of the fixed point we give. Provided that we don’t identify “having semantic value 1” with “being true” we have two possibilities: either the designated value is only “true” and we obtained KFS, the theory which Field’s one is based on; or we take as designated all values higher than 0, namely both “true” and the third value – in the specific case, “both true and false” – and in this case we obtain the dialethic logic LP. As far as the differences between the two theories are concerned, they help us to highlight the duality between them. In fact, both theories face the Curry’s paradox and the solutions they provide mirror each other: Beall rejects the rule of the elimination of the conditional (known as *modus ponens*), whereas Field rejects its counterpart, namely the rule of the introduction of the conditional. Also from a pragmatic point of view, the two theories don’t disprove the thesis that they are dual. In fact, both Beall and Field reject the interdefinability of the notions of acceptance and rejection. The reasons adduced by the two philosophers are different but complementary to one another. According to Beall, there are situations where we are allowed to accept both a sentence and its negation, whereas, according to Field, it’s sometimes possible to reject both a sentence and its negation. However, both philosophers maintain the exclusivity – that is, the incompatibility – of the two notions. Lastly, if we look at the problems faced by the two theories, things are not so different. Both dialetheists and paracompletists are facing a parting of the ways: on the one hand, they can solve the revenge liar only providing that they abandon the completeness of their languages; on the other hand, the second possibility is to keep the language expressively complete, on pain of insoluble contradictions.

Therefore, the duality between the two theories that clearly appears from the analysis leads to the conclusion that we can in no way establish which one is the best theory, because neither of them has some decisive advantages or disadvantages over the other. If the two theories are good, they are equally good. To decide which one is the best means taking an arbitrary choice only on the basis of the following observations: on the one hand, we have a theory that keeps consistency as a bounding requirement, but, as a consequence, has the unpleasant result of a very complicated and entangled construction aimed towards an adequate conditional; on the other hand, we have a theory that has simplicity as a distinctive feature of its structure, but that, in contrast, must admit inconsistency among its main theses. Therefore, the choice is between consistency – for Field’s paracomplete theory – and simplicity – for Priest and Beall’s dialetheism – and it cannot be based on nothing but an arbitrary preference on these aspects, because, as we saw, from a logico-semantic point of view there is nothing else that allows us to choose between the two theories.

In conclusion, this fact is the reason why the original question receives a negative answer, that is, actually deflationists should not be dialetheists, contrariwise to what Armour-Garb and Beall claim. The reason is not that dialetheism is not an adequate theory in order to respect the deflationary intuitions about the nature of truth, but rather that it is not the only available theory that allows deflationists to semantically express their ideas – also Field’s paracomplete theory can do this job. Moreover, these two theories – dialetheism and Field’s paracomplete theory – have the same advantages or

Do Dialetheism and Deflationism Need Each Other?

disadvantages. Hence, this section suggests the conclusion that deflationists can arbitrarily choose between a dialetheic theory and Field's paracomplete one.

Chapter 5

Alethic Fictionalism, Aboutness and The Liar

5.1 Constructive Methodological Deflationism

5.1.1 Introductory Remarks

In chapter 4, I analysed the consequences of the possible interactions between dialetheism and a deflationary theory of truth. In particular, I tried to provide an answer to two fundamental and complementary questions on this matter, that is “Should dialetheists be deflationists?” and “Should deflationists really be dialetheists?” Both questions had a negative answer. However, to say that deflationists don’t need to assume a dialethic theory and that dialetheists don’t need to embrace a deflationary theory of truth, doesn’t mean that the two theories cannot get along with each other, but simply that the union of the two theories is not the only choice for them. In other words, on the one hand, there is no evident benefit for the deflationist in subscribing a dialethic account because there is at least one other equally good theory available to them, i.e. the paracomplete theory developed by Field; and similarly, on the other hand, there is no evident benefit also for the dialetheist in embracing a deflationary theory of the nature of truth because the Dummetian theory developed by Priest provides the same advantages and disadvantages of Beall’s one for a deflationist.

In this chapter I’m going to bring alethic fictionalism into the debate, that is, the last perspective about the notion of truth that so far has been a peripheral issue in the analyses. I’m referring to the perspective about truth-talk or, better, the way we interpret, or should interpret, the truth-talk. Hence, focus of the present chapter is a particular perspective of truth-talk, namely, truth-theoretic fictionalism. The aim of the present chapter is twofold: on the one hand, I wish to figure out whether fictionalism can get along with dialetheism and deflationism, or not; and, on the other hand, I’m going to present the new approach about meaning recently developed by Stephen Yablo.¹

The chapter is inspired by a paper signed by J.C. Beall,² in which the philosopher wishes to keep all the three theories about truth we analysed together: *dialetheism* that, despite being a metaphysical theory, deals with the logico-semantics structure of the truth predicate and the rules regulating its use

¹ Yablo[2014].

² Beall[2004].

within our language; *deflationism* that deals with the ontological aspects of the truth predicate, meaning that deflationists wish to tell us what the real nature of truth is;³ and *alethic fictionalism* that, as we just said, is a theory about the discourse of truth, that is, an alethic fictionalist tells us how we interpret, or should interpret, the region of discourse involving the truth predicate.

Starting from this idea, the purpose of my analysis, in this chapter as well as in the previous one, is to answer two complementary questions. In the first place, I'll wonder whether the dialetheist should be fictionalist – that is, whether there is some advantage for the dialetheist in assuming a fictionalist perspective of truth-talk – or not. In order to answer this question I'm going to compare dialetheism with those inconsistency theories of truth that need alethic fictionalism in order to provide a motivation for their account. In other words, I'll wonder if the reasons moving inconsistency theorists towards the endorsement of a fictionalist perspective of truth-talk can be shared also by the dialetheist.

In the second place, I'll wonder the complementary question, that is, whether the fictionalist really should be dialetheist, as Beall claims, or not. The remarks and considerations made by David Liggins on this matter will come to my aid in the attempt to respond to this latter question.⁴

Finally, since the analysis accomplished so far will have the result that the attempt made by Beall to construct a theory keeping deflationism, fictionalism and dialetheism together does not work, I'll present an alternative account that, as I wish to prove, can be wholly translated into a fictionalist account. In fact, after a detailed overview on the theory of subject matter and partial truth developed by Yablo, I'm going to show that this theory can be translated into a fictionalist theory even if Yablo never makes this possibility explicit inside his book. Furthermore, I'll try to show that it is possible to apply Yablo's account of subject matter also to truth and, in particular, to semantic paradoxes, like the liar, in order to provide an alternative solution to them, sharply different from the existent ones.

5.1.2 Beall's Account

As we have stressed again and again in chapter 4, Beall develops a dialethic theory capable of validating both side of the T-schema – like Priest's theory does – and the intersubstitutivity principle – contrariwise to Priest. In this way, Beall obtains the full transparency of the truth predicate that is exactly what makes Beall's theory a deflationary theory of truth. Very briefly, this is what we know about Beall's theory. But, let's take a step back and focus on a previous idea by Beall. In a paper published in 2004,⁵ Beall develops a theory he calls *Constructive methodological deflationism*, namely, an account of truth that has all the right features to be labelled as a fictionalist approach of the discourse about truth.⁶

³ Cfr. Introduction, footnote 2.

⁴ Liggins[2014].

⁵ Beall[2004].

⁶ It's worth mentioning that this is not Beall's last work on this topic. As I stressed again and again in chapter 4, Beall has changed his mind in his more recent works, such as Beall[2009].

The first step of the argument in favour of constructive methodological deflationism is a mental experiment. Beall asks us to imagine of removing the truth predicate from our natural language, i.e. English. In this way we can't use formulas like "is true", "is false" and the like anymore. In place of truth and falsity predicates, the language is provided with two other predicates: "Aiehtela accepts" and "Aiehtelanu accepts". Now, all we have to do is to teach children the functioning of these two new predicates. In order to do it, we use schemata specifying the behaviour of the new predicates, because they are easier to understand and remember for them. The most important schemata are the following:

- (A1) Aiehteia accepts "A" iff A.
- (A2) Aiehtelanu accepts "A" iff $\neg A$.
- (A3) Aiehtela accepts " $A \wedge B$ " iff Aiehtela accepts "A" and Aiehtela accepts "B".
- (A4) Aiehtelanu accepts " $A \wedge B$ " iff Aiehtela accepts " $\neg A$ " or Aiehtela accepts " $\neg B$ ".
- (A5) Aiehtela accepts "A" or accepts " $\neg A$ " (and the same goes for Aiehtelanu).

for every A, where A is a declarative sentence of the language.⁷

The last schema, i.e. (A5), is very important because it holds that the logic behind the predicate "Aiehtela (Aiehtelanu) accepts" is exhaustive, namely, a logic that doesn't admit indeterminacy.

More in details, Beall uses LP as a base logic regulating the functioning of the two new predicates.⁸ For this reason, it very clearly appears that an inconsistent behaviour characterizes the new predicates. It seems, in fact, perfectly plausible for Aiehtela (or, Aiehtelanu) to accept both a sentence and its negation. However, there are other problematic cases we should take into account. I'm referring, for instance, to those situations where Aiehtela seems to both accept and not accept the same sentence at the same time under the same circumstances. How should we behave towards those cases? How do we have to consider them? The quick and immediate reaction to them will be of difficult understanding. It seems that children have some difficulties in picturing and, hence, to understand what those situations mean and how it is possible for them to happen. However, if we take into account some special sentences, then those situations become clearer. For instance, let's consider the following sentence:

(*) Aiehtela does not accept the starred sentence.

⁷ Beall[2004], p.201.

⁸ If we refer to Priest's logic, it is clear that we won't have the following principle, as we stressed again and again:

- (A6) If Aiehtelanu accepts "A", then Aiehtela does not accept "A".

On the contrary, (A6) holds in Beall's theory. In any case, whether we assume the principle or not makes no difference for Beall's argument.

From what the starred sentence says and by the principles governing the functioning of the predicate “Aiehtela accepts” – i.e. A1, A2 and A5 – follows that Aiehtela both accepts and does not accept the starred sentence.⁹

Despite the children being initially confused from this analysis of the starred sentence, then, after much consideration, they come to the following conclusion:

Given the “semantics” (the basic rules governing the relevant terms), the resulting inconsistency is likewise inevitable. [Beall 2004, p.203]

Hence, after careful reflection upon the starred sentence, the children manage to picture out the situation and, in particular, they manage to understand that the situation displayed by the starred sentence correspond exactly to the situation where Aiehtela and Aiehtelanu accepts the same sentence. The fact that the situation can be described also as the conjunction of the starred sentence with its negation is only a consequence of the logic we are using, i.e. LP, and of the specific story. In other words, the fact that we can describe the situation as the conjunction between the starred sentence and its negation is simply a *spandrel* – to use Beall’s well-known vocabulary – of the new predicates “Aiehtela accepts” and “Aiehtelanu accepts”.

Granted that the children, in fact, understood the logico-semantic part of the functioning of the two new predicates, they may wonder now upon metaphysical issues about those predicates. In particular, they may wonder what Aiehtela and Aiehtelanu are, i.e. what the nature of the two predicates is. However, according to Beall, this is not a problem because we are prepared to provide an answer to their ontological questions: we may say that Aiehtela and Aiehtelanu are nothing but expressive devices helping us in the expression of generalizations we cannot make otherwise within our language.¹⁰ For instance, without Aiehtela we couldn’t say things like the following:

(A) Aiehtela accepts everything the Pope says.

and, similarly, for Aiehtelanu. The ontological question, hence, disappears because there is no underlying nature of Aiehtela and Aiehtelanu to investigate.

The fundamental point of Beall’s argument is that truth behaves exactly in the same way as Aiehtela. The linguistic role of the two predicates is the same:

⁹ Another example that is worth mentioning and that uses the other predicate, i.e. “Aiehtelanu accepts”, is the following:

(**) Aiehtelanu accepts the double starred sentence.

Also this sentence displays a problem, because from (**), among with the principles regulating the functioning of “Aiehtelanu accepts”, follows that Aiehtelanu both accepts and does not accepts the double starred sentence.

¹⁰ Remember that this feature of the truth predicate is not accounted only by a deflationary theory but, on the contrary, it is accepted almost by everyone, whatever theory of the nature of truth one decides to embrace, whether she is a deflationist, a Dummetian, an inflationist or any other sort of theoretician about the nature of truth. What characterizes the deflationary perspective from the others is the claim that everything we can say about truth is limited to its functioning as an expressive device. According to deflationism, what we can say about truth is limited to the instances of the T-schema because it has no underlying nature that goes beyond that schema.

“is true” has the same linguistic role of “Aiehtela accepts” and “is false” has the same linguistic role of “Aiehtelanu accepts”. Moreover, there is no difference between the story of Aiehtela and Aiehtelanu and the story of truth and falsity. In addition, there is no difference also between the discourse about Aiehtela and truth-talk.¹¹ The two discourses enable us to perform the same function, namely, to express generalizations we cannot express otherwise.

Thus, as Aiehtela very clearly appears to be a construction – i.e. a device of semantic descent introduced within the language in order to perform an expressive role –, the same goes for truth, by analogy. In fact, when Beall talks of “construction”, he means what follows:

A construction is a story that we make up (that we construct), usually for some given purpose. Some construction may be true; some false. Any fiction is a construction; however, the converse isn't true, since (as the term is generally used) no fiction is true. [Beall 2004, p.208]

If we don't realize the similarities between the two stories – that is, if we don't realize that also truth, as well as Aiehtela, is merely a linguistic construction – is only because we have used it ever since. On the contrary, Aiehtela is a new concept and, for this reason, we find easier to accept it as a construction. Moreover, when we use “Aiehtela accepts” we are intentionally engaged in *as if* thinking – that is, a game of make-believe – and the same exactly goes for “is true” as well. In both cases we pretend that these things – respectively, Aiehtela and truth – exist merely for expressive reasons. We can sum up this idea with Beall's words:

We speak as if we are describing things as having or lacking properties called “truth” and “falsity”, in order to make claims of other (more complicated) sorts indirectly. [Armour-Garb and Woodbridge 2015b, p.130]

and the same goes for Aiehtela. Hence, even if we substitute truth-talk with talk about Aiehtela, or vice versa, nothing will change at linguistic level because their role within the language is exactly the same.

The meaning of the name of Beall's theory – i.e. constructive methodological deflationism – becomes, now, very clear. *Constructive* pinpoints the characteristic feature that truth assumes in Beall's theory, namely, that it is nothing but a construction, in the same way as the new predicate, “Aiehtela accepts.” This is where fictionalism comes in. We make *as if* truth exists but it actually is only a construction, a fiction. *Methodological* refers to Beall's claim already mentioned that issues about the nature of truth can be set aside for the prosecution of his current analysis.¹² As Beall puts it:

¹¹ In fact, Beall explains that the analogies persist if we let the issue about the nature of the two predicates aside. According to Beall, this decision about the nature of truth doesn't cause any problem for his argument. The reasons appeal to Ockham's razor, according to which if two hypotheses are equally good in their explanation of the data, then we must choose the simplest. The deflationary approach wholly explains the functioning and the role of the truth predicate in the language. Therefore, there is no need to postulate a robust nature of truth. For the details see Beall[2004], p.206.

¹² cfr. footnote 11.

It strikes me as a sound methodological principle to address such questions [i.e. about the nature of truth] only if there is strong reason to do so. [Beall 2004, p.207]

Beall in this quote is referring to Field, who maintains that the starting point of our analysis should be deflationism, because it is the most economic theory from an ontological point of view. We should embrace a more serious ontological theory only if deflationism turned out to be untenable. Until this very moment, we can set the issue about the nature of truth aside.¹³ Lastly, the theory deserves the label of *deflationism* because it satisfies all the criteria imposed by a deflationary theoretician. Constructive methodological deflationism is, to all intents and purposes, a form of deflationism because it maintains that the truth predicate has no underlying nature and that everything we can say about truth is limited to the instances of the T-schema and to the fact that truth is nothing but a device of semantic descent.¹⁴

5.2 Should Dialetheists Be Fictionalists about Truth-Talk?

As we saw, the argument presented in the previous section is the attempt to keep all the three different perspectives about the truth predicate we have taken into account – namely, dialetheism, deflationism and fictionalism – together. In chapter 4 I analysed in details the consequences of the possible interactions between the first two perspectives, i.e. dialetheism and deflationism, and, in particular, I tried to figure out if there is an effective advantage for the dialetheist in the assumption of a deflationary account of truth, and vice versa. The focus of this chapter is the third perspective, i.e. fictionalism about truth-talk. In this section I'll wonder whether the dialetheist takes an advantage of a sort in the assumption of a fictionalist perspective of truth-talk, or not. In other words, I'll wonder whether alethic fictionalism can be in some way useful for the dialetheist, or not. More precisely, I'll wonder whether fictionalism about truth-talk is useful in a way forcing the dialetheist to embrace it instead of another approach of truth-talk, or not. More directly, the

¹³ Field[1994].

¹⁴ The theory, as I presented it, does not validate the intersubstitutivity principle, that, remember, is a peculiar feature for a deflationary theory of truth in its common interpretation. Beall provides a twofold reason why this does not end in a problem: on the one hand, what is sufficient for a theory in order to be deflationary is that “all relevant uses of ‘is true’ be explained in terms of the basic disquotational features of ‘is true’” [Beall 2004, footnote 26] and this is exactly what happens in Aiehtela's story; on the other hand, if this still didn't satisfy deflationists, then there would be no problem in supplying the theory with a principle of intersubstitutivity, that is, there is nothing in the theory that would prevent the introduction of such a principle. The only problem that, perhaps, could arise is that if the intersubstitutivity principle were valid, a “friendly” paraphrases of the contradiction due to the conjunction of the starred sentence and its negation would not be possible anymore. This does not appear to disturb Beall in any way because, according to him, the constructive character of the discourse about Aiehtela is what allows us to arbitrarily decide how to describe the situation arising from the starred sentence.

question I wish to answer here is the following: should dialetheists be truth-theoretic fictionalists?

The section will proceed through the comparison between dialetheism and inconsistency theories of truth. The reason of this choice is that many among the advocates of the latter maintain that alethic fictionalism is a methodological necessity for their theories because it allows them to justify certain specific claims about truth and its use. After having briefly make an overview of the main characters of inconsistency theories of truth, I'll wonder whether the reasons moving the advocates of the latter in the adoption of a fictionalist perspective of truth-talk can be found also within dialetheism, or not. Do dialetheists share this necessity with advocates of inconsistency theories of truth, or, on the contrary, they don't face the same problems? The conclusion of this analysis will allow me to affirmatively or negatively answer the main question of the section.

5.2.1 The Inconsistency Theories of Truth

The philosopher who is mostly associated with inconsistency theories of truth, on the heels of Alfred Tarski, is Matti Eklund, who, in a series of papers, defends the idea that the language is inconsistent.¹⁵ A more accurate way to define an inconsistency theory of truth is the following:

The inconsistency theory of truth can then be formulated as the view that the rules governing the appropriate use of the truth predicate are inconsistent, in the sense that they compel us to accept both a sentence and its negation. [Burgess 2007, p.13]

Therefore, the idea on the root of inconsistency theories is that the principles governing the truth predicate are, in fact, inconsistent.

The idea of an inconsistency theory of truth didn't meet a lot of interest among philosophers of truth. Among the few names that, besides Eklund, are worth mentioning there are John P. Burgess and his son Alexis Burgess,¹⁶ Jody Azzouni,¹⁷ Charles Chihara,¹⁸ Kirk Ludwig¹⁹ and, of course, Alfred Tarski.²⁰ The most interesting account for my analysis is the conjoint work of J.P. Burgess and his son A. Burgess and the work made by A. Burgess alone in his doctoral thesis. Burgess and Burgess defend an inconsistency account of truth, but they take inconsistency theories a step forward by arguing that they need to adopt a fictionalist perspective of truth-talk. We will see in next section this last argument and the reason behind it. I'm going to focus here mostly on the main features of the theory developed by Eklund, in order to provide a brief overview of inconsistency theories.

In a nutshell, Burgess and Burgess explain what is the idea behind Eklund's theory:

¹⁵ Eklund[2002a, 2002b, 2005 and 2008].

¹⁶ Burgess[2007] and Burgess and Burgess[2011].

¹⁷ Azzouni[2003].

¹⁸ Chihara[1979].

¹⁹ Ludwig[2002].

²⁰ Tarski[1983].

A first answer is suggested by Matti Eklund's endorsement of the principle that when nothing can meet all the conditions built into the meaning of some term, then the distinction the term in fact marks (or should be understood as marking) is the one that comes *closest* in doing so. [Burgess and Burgess 2011, p.240 (emphasis in original)]

Let's see, now, in details what this means.

The starting point of Eklund analysis is, like for most philosophers and theories of truth, the liar paradox.²¹ As we stressed again and again, the liar is a sentence that says of itself that it is false. We have formalized it in the following way:

(L) (L) is false.

It is worth remembering the analysis of the liar sentence. Let's go by cases. If we assume that the liar is true, then, for what it says, it turns out to be false. Hence, we have an implicit contradiction. Let's suppose, now, that (L) is false. This is exactly what the liar says, so (L) is true. Hence, from these considerations another implicit contradiction follows. Therefore, the liar sentence leads to contradiction whatever the valuation we try to assign it is.

The most common way to solve the problem is trying to find out where the mistake in this argument is. If an argument leads to contradiction, then there must be some sort of mistake somewhere in the reasoning. This is what commonly philosophers and logicians think. Therefore, traditionally, people try to reject one or more principles governing our language – and, in particular, governing the functioning of the truth predicate within our language – and that are accused to be the cause of the contradiction. In particular, if we focus on classical logic, the choice classical logicians make in order to solve the liar is to restrict the T-schema, either by rejecting one side or the other of the schema, or by claiming that not all its instances hold and excluding, in particular, those instances that leads to contradiction, such as the instance of the liar.

Differently from these approaches, Eklund maintains that the real reason why semantic paradoxes arise is the intrinsic inconsistency of the expressions involved in them. Eklund is inspired by Tarskian intuition that what semantic paradoxes do is to highlight the intrinsically inconsistent character of our natural language.²² More in details, Tarski explicitly claims what follows:

A characteristic feature of colloquial language (in contrast to various scientific languages) is its universality. [...] If we are to maintain this universality of everyday language in connexion with semantical investigations, we must, to be consistent, admit into the language, in addition to its sentences and expressions, also the names of these sentences and expressions, and sentences containing these names, as well

²¹ Eklund focuses also on the *sorites paradox*. Take a sorites series of coloured patches ranging from red to orange. The premise from which the reasoning starts is that for every patch K in the series, if K is red, then K+1 is red too. The first patch looks red and the last looks orange. By stipulation, if the first patch looks red, the second patch, that is adjacent to the first, looks red too. But, we can conclude that all the patches are red because the same reasoning must be applied at every step. Hence, the last patch looks red. But, we know that the last patch looks orange, not red. Hence, we have a paradox.

²² Tarski[1983].

as such semantic expressions as “true sentence”, “name”, “denote”, etc. But it is presumably just this universality of everyday language which is the primary source of all semantical antinomies, like the antinomies of the liar [...] These antinomies seem to provide a proof that every language which is universal in the above sense, and for which the normal laws of logic hold, must be inconsistent. [Tarski 1983, p.164]

And a language L is inconsistent if and only if the principles constitutive of the meaning of its expressions are jointly inconsistent, where a principle is constitutive of the meaning of an expression “just in case competence with the expression involves being disposed to accept the principle as true or valid.”²³

Eklund’s intuition is that the speakers are willing to accept those principles that are accused to be the cause of the paradoxical conclusion – i.e. the full T-schema without restrictions, that is, with both its side and its instances – in virtue of their semantic competence. In other words, our semantic competence with the truth predicate is the disposition to believe any instance of the T-schema. More in details, Eklund’s idea is that the liar arises because our semantic intuitions are, in fact, inconsistent, and by semantic intuitions we mean “the intuitions of truth and validity that we have by virtue of our semantic competence.”²⁴ The problem with the liar is that we are disposed to accept the invalid steps in the liar argument because of our semantic competence with the expression involved. Therefore, Eklund agrees that there really are one or more invalid steps in the reasoning of the liar leading to the contradictory conclusion. What he criticises of classical approaches is that it is not enough to say of a specific step that it is unsound and that a specific inference is not correct, because we must also explain why that inference seemed to us true and valid. Otherwise, according to Eklund, we would be accused of “manifest lack of semantic competence.”²⁵

The main difference, thus, between Eklund’s inconsistency theory and the others is that, according to Eklund it is indeterminate which are the untrue steps in the liar reasoning and, more importantly, we are inclined to accept those wrong steps because of our semantic competence, and, in particular, because there is no “defeating evidence against its truth or validity.”²⁶ In the words of Eklund:

What is special about my view is just that I hold that the semantic values of these expressions do not actually, on pain of contradiction, make the meaning-constitutive principles for these expressions true: but are only what come closest to making these meaning-constitutive principles true. [Eklund 2002a, p.323]

Where for *meaning-constitutive principles* he means those principles that are accepted by anyone is fully competent with the expressions involved in them. Eklund makes a difference between the concept of being meaning-constitutive and the concept of being analytic for a principle, where analytic principles are the principles that are true by virtue of meaning. Moreover, when the meaning-

²³ Eklund[2002b], p.256.

²⁴ Ibid. p.253.

²⁵ Ibid. p.252.

²⁶ Ibid. p.252.

constitutive principles are inconsistent, Eklund claims that “the semantic values of the expressions in the language-fragment are what come closest to making the meaning-constitutive principles true.”²⁷ Eklund adds another concept: a sentence *exerts pull* if and only if the speakers are disposed to accept it by virtue of their semantic competence and, as we will see, Eklund’s aim is to prove that the liar, in fact, exerts pull.²⁸

It is also necessary to specify which is the notion of inconsistency Eklund uses in his account, which, it is now clear, is very different from the traditional meaning of inconsistency. Standard inconsistency is what Eklund calls here “Priest-inconsistency”. A language L is *Priest-inconsistent* just in case the semantic of negation, “ \neg ”, is such that it allows that there is at least a sentence α of L such that both α and $\neg\alpha$ are true.²⁹ Eklund defines a notion of deep inconsistency, according to which, “a language L is *deeply inconsistent* just in case the semantic values of expressions of L cannot be such as to make all principles constitutive of the meanings of these expressions true and valid, respectively.”³⁰ Hence, a language can be deep inconsistent without also being Priest-inconsistent. In fact, for Priest-inconsistency is not sufficient that “the meanings of the expressions are such that the principles leading up to the contradiction are demanded by the meanings to be true.”³¹ Therefore, Priest-inconsistency is a notion stronger than deep inconsistency.

These are the main key elements of the theory developed by Eklund. Let’s now see how they can be concretely applied within the theory. In particular, with the background just specified, we can use Eklund’s very words to explain which is the main point of his inconsistency theory:

What I think this shows is that we must not assume that if the meaning of “true” requires that the disquotation schema is valid, then the extension of “true” is such that it is valid. A different, more plausible picture is available: the meaning of “true” requiring that the disquotation schema be valid only implies that a disposition to accept the disquotation schema is part of semantic competence with the predicate. [Eklund 2002b, p.256]

Hence, the idea is that there are certain expressions in the language the main comprehension of which involves our disposal to accept certain false claims about them.³²

²⁷ Eklund[2002a], p.322.

²⁸ Ibid., p.322.

²⁹ Ibid., p.324.

³⁰ Ibid. (emphasis in original).

³¹ Ibid.

³² Someone could think that, as a consequence, semantically competent speakers are disposed to accept anything. Eklund emphasises that this is not the case. The objection goes as follows: semantically competent speakers are disposed to accept certain instances of the T-schema and, at the same time, because of their very same semantically competence, they are disposed to accept also the paradoxical arguments that have those instances of the T-schema as premises. The problem, thus, is that from Eklund’s idea follows that semantically competent speakers are disposed to “accept inconsistent verdicts also on sentences whose truth-conditions seem unproblematic.” [Eklund 2002b, p.266]. According to Eklund, semantically competent speakers have a grasp of how assign truth-values to sentences of the language in such a way that they are able to distinguish between pathological sentences and regular sentences.

From Eklund's point of view, the difference between meaning-constitutive and analytic principles allows us to say of the principles governing the liar reasoning that they are meaning-constitutive, but not analytic, because they are not true in virtue of their meaning. This is exactly the reason why Eklund says that the principles regulating the truth predicate are inconsistent but not true and this, as we have just saw, is exactly what Eklund calls deep inconsistency.

Therefore, it is now clear why according to Eklund the language is inconsistent. According to Eklund's inconsistency theory, one of the steps of the liar reasoning, in fact, fails – exactly like standard accounts maintain –, but which one is indeterminate. If we assume classical logic, we must reject one principle between T-Elimination and T-Introduction. The problem, however, is that in order to determine exactly which of the two rules fails, we should determine first whether the distinction between true and false that “comes closest to meeting the conditions built into the meaning of the truth predicate is one that marks the liar paradox as true or one that marks it as untrue”³³ – or false in the case of the standard liar, L. The problem with this distinction is that it is far from easy task. Eklund maintains that an account of truth is correct only if it comes closest to making the meaning-constitutive principles true. Hence, even if we understand “comes closest to” as “validate as many instances of the T-schema as possible” – that is, as the account that maximally preserves the meaning constitutive principles of truth –, that won't solve the problem because there are several maximal consistent sets of instances of the T-schema.³⁴

To make things easy, it is worth remembering that, according to Eklund's inconsistency theory, at least one of the steps involved in the liar reasoning is invalid, but by our semantic competence we are disposed to accept it. This, by definition, means that the liar *exerts pull*. For this reason, Eklund says that also the invalid steps in the argument are part of the meaning of the liar. Therefore, this is exactly the reason why Eklund can say of the meaning-constitutive principles of the language that they are, in fact, inconsistent and, more specifically, that they are inconsistent in Eklund's meaning of inconsistency, that is, they are deep inconsistent.

Hence, to conclude, Eklund's idea can be summed up in the following way:

Natural languages are semantically rich but logically non-trivial and on which speakers' coherent use of them depends on their knowledge (“grasp”) of how they manage to be so. This common view is supplemented by claims to the effect that semantic competence among speakers involves a disposition to accept inconsistent “meaning constitutive principles” and hence paradoxical arguments, but this adds little to what anyone would grant: that speakers who seem to understand “is true” in a natural language as well as anyone does find the paradoxical arguments to involve premises that look true and inferences that look

³³ Burgess and Burgess[2011], pp.240-241.

³⁴ Moreover, as Burgess and Burgess notice, we don't even have the possibility to decide which T-biconditionals we should validate on the basis of a putative priority of some kind, because “the equivalence principle is simply a universal generalization, and as such does not discriminate among its instances.” [Burgess and Burgess 2011, pp.241-242]

valid, but to have conclusions that look false. Eklund adds to this only insistence that being disposed to accept the premises and inferences is a condition of full competence. Eklund thus [...] has to accommodate the evidence as to what speakers of natural languages take them [i.e. paradoxes] to be like while at the same time showing how the language speakers speak admits of a semantic treatment on which it is expressively rich but logically non-trivial. [Patterson 2009, pp.409-410]

There are other consistent alternatives to inconsistency theories of truth that we can find in the literature about the concept of truth and, in particular, about the liar paradox. It is worth mentioning two well-known theories among the others: the construction of the fixed point developed by Kripke³⁵ and the paracomplete theory developed by Field.³⁶ However, according to Burgess, these two theories have two big disadvantages that inconsistency theories don't have: on the one hand, they apply only to semantically impoverished languages; and, on the other hand, even if they could be developed in a way that cover also natural languages, they would be too complex if compared with inconsistency theories anyway, especially they would be too complex to be understood from everyone, included the ordinary speaker.³⁷

5.2.2 Inconsistency Theories and Alethic Fictionalism

So far we have specified the main characters of an inconsistency theory of truth. What we didn't take into account yet is the kind of attitude we must have towards sentences belonging to truth-talk if we accept an inconsistency theory of truth. As we have already mentioned, Alexis Burgess, in his doctoral dissertation,³⁸ maintains that truth-theoretic fictionalism is the best option for an advocate of inconsistency. In section 3.4 we defined truth-theoretic fictionalism as the idea that the discourse about truth must not be understood literally, but rather as a convenient fiction.

In the first place, we must mention that, according to Burgess, fictionalism about truth-talk is justified by an apparent inevitability of the revenge phenomenon of semantic paradoxes. Remember the functioning of the revenge phenomenon with regard to the liar paradox: every new attempt to solve the liar paradox leads to diagnose some lack in a specific class of sentences; the vocabulary used to formulate this diagnosis can be used, in turn, to formulate a new version of the liar paradox that, for this reason, is called revenge (or strengthened) liar paradox. Every new solution to the paradox gives rise to a new more powerful paradox that cannot be solved using the same key notions as before, and every theory faces its own revenge paradox. The most known version of strengthened liar is the following:

(Q) (Q) is not true.

³⁵ Kripke[1975].

³⁶ Field[2008].

³⁷ Beall[2004], footnote 21.

³⁸ Burgess[2007].

The particular aspect of this formulation is that it does not involve falsity, but rather non-truth. In particular, (Q) is a strengthened version of the liar paradox for all those theories that postulate the existence of gaps among truth-values, namely, those theories claiming the liar to be neither true nor false, i.e. to have an indeterminate truth-value.

According to Burgess, the problem taken up by the revenge phenomenon of the liar does nothing but emphasise the intrinsically inconsistent character of the language. As we said in the previous section, this is the starting point of inconsistency theories of truth that, remember, have the following purpose:

The so-called inconsistency theory of truth [...] purports to explain this tenacity of semantic paradoxes by maintaining that some suitably unrestricted version of the T scheme is indeed “built in” to our very concept of truth. [Burgee, forthcoming]

Therefore, the advocates of inconsistency theories of truth maintain the thesis that the language is intrinsically inconsistent because this is the only explanation we can provide for the semantic paradoxes – especially of their revenge versions – and, at the same time, they preserve the very intuitive idea that the full T-schema is a fundamental element of the concept of truth, that, in other words, means that the T-schema gives us – and, perhaps, exhausts – the whole meaning of the truth predicate.

The advocates of inconsistency, however, don't want to take this point a step forward by ending up in the idea that true contradictions really exist. According to Burgess, the most adequate “sensible morals to draw from the inconsistency theory”³⁹ is to adopt a nominalist or agnostic theory about the nature of truth. The question now is: what do we have to do with this kind of problematic and inconsistent concept such as truth? Do we have to reject it or to keep using it? On the one hand, ever since the attitude towards a defective concept has been of rejection and sometimes it leads to the replacement of that concept with a new unproblematic one, by completely revising the region of discourse about it. On the other hand, however, in the specific case of truth, inconsistency seems not to present a big pragmatic deal because, as also Priest maintains, consistency is the case by default and, thus, it's very difficult for paradoxical situation to appear in ordinary conversations. This is why advocates of inconsistency don't want to abandon a concept, such as truth, because of its inconsistency. If this inconsistency is a rarity and it appears only in certain specific domains, then there is no reason to completely abandon the concept of truth. Anyway, despite that, they face a dilemma, because, of course, knowing that inconsistency can appear is not reassuring.

What should we have to do now? What do we have to choose? According to Burgess, truth-theoretic fictionalism is the perfect alternative because it takes place halfway between the two corners of the dilemma, the one who wishes to reject the concept of truth as we know it and the one who wishes to keep using it unchanged. The kind of fictionalism Burgess has in mind can be both descriptive and prescriptive, that is, it can be both a hermeneutic and a revolutionary kind of fictionalism. Burgess's idea can be summarized in the following way:

³⁹ Burgess[forthcoming].

Chapter 5

We can keep using the truth predicate, but use it “lightly” from here on out, as a figure of speech or convenient fiction. Even if nothing is true – even if nominalism about truth [...] is the right conclusion to draw from the inconsistency theory – there should be no obstacle to our talking and thinking as though many things were. [Burgess, forthcoming]

As a fictionalist approach, Burgess approach tells us that when we utter a sentence involving the truth predicate, we don't have to take it literally, but rather we have to interpret the utterance in a fictional way, that is, we have to interpret the utterance as belonging to the relevant fiction, i.e. the fiction of truth.

Within the fiction, the truth predicate obeys the full T-schema in exactly the same way inconsistency theories suggest. Despite the inconsistent character of truth within the fiction, this does not lead to the conclusion that we must accept true contradictions. The two things are – and must be taken – separated. Burgess himself says:

Engaging with an inconsistent fiction doesn't involve any commitment to contradictory propositions. [Burgess, forthcoming]

The reason is that all inconsistency theories embracing a fictionalist perspective of truth-talk are committed to is that within the fiction of truth there are true contradictions. Alternatively, what we are allowed to say is only that according to the fiction of truth there are true contradictions. However, this latter claim, obviously, doesn't commit one to the existence of true contradictions *tout court*, that is, the existence of dialetheias outside the fiction as well. Moreover, whoever is engaged in the fiction of truth does not really assert and believe true contradictions, but rather she only pretends to assert and believe them, that is, she merely makes *as if* she asserts and believes that there are true contradictions.

The first thing that comes into our mind when we speak of fictionalism about truth-talk is that there is something wrong, something not completely convincing in this idea, because defining an inconsistency theory of truth as a fiction seems to be something false. However, there is a remark it is worth doing here. On the one hand, from a fictionalist perspective, this corresponds to saying that according to the fiction of the inconsistency theory of truth, the inconsistency theory of truth is false; and, on the other hand, it seems perfectly legit for someone to believe that according to the fiction of the inconsistency theory of truth, the inconsistency theory of truth is true. This might appear very strange to someone. However, this is not a problem from an inconsistency theory's point of view. As a matter of fact, from the point of view of an advocates of inconsistency theories, there's nothing strange or wrong in the claim that something is true and false, because this assertion can be made exclusively within the relevant fiction, that is, only within the fiction of truth, and it is not meant to be an assertion about the real world and, thus, it is not committing us to the existence of true contradictions outside the fiction.

Moreover there is one more remark that is worth mentioning: from the fact that something is both true and false inside the fiction does not automatically follow that everything is true inside it. In other words, a contradiction in the fiction doesn't trivialize the fiction itself. As a consequence,

the inconsistency theory of truth is not trivial. The reason is that the inconsistency theory of truth is not properly a theory in the true sense of the word but, as a fiction, it can be imagined as a story including everything analytic to the notion of truth.⁴⁰ To restate this idea in a more explicit way:

Inconsistency doesn't necessarily trivialize stories in the same way that it trivialized theories. [Burgess 2007, p.37]

According to Burgess, another decisive advantage of the assumption of a fictionalist approach of truth-talk for an inconsistency theory is that this perspective can accommodate also the correspondentist intuitions about the nature of truth, that is, it can involve also the claim that something is true just in case it corresponds to reality, claim that, because of this very reason, is taken as analytic of the notion of truth. Moreover, this move does not force the alethic fictionalist to provide and develop a theory of reference because fictional stories are notoriously incomplete, that is, certain sentences are neither true nor false according to them. Therefore, the advantage of a fictionalist approach of truth-talk is that, somehow, it provides a peaceful solution to the dispute between deflationists and other realist advocates about the nature of truth.⁴¹

5.2.3 Dialetheism and Alethic Fictionalism

As we stressed again and again, the great advantage in the endorsement of a fictionalist perspective of truth-talk by an advocate of inconsistency theories of truth is that fictionalism allows her to solve the main problem faced by all theories of truth, namely the liar, and, at the same time, allows her not to relapse on a dialethic theory, according to which there are true contradictions also outside the fiction.

According to alethic fictionalists, the T-schema is analytic of the notion of truth that, by definition, means that when we use the truth predicate with its common meaning, we, in fact, are committed to the acceptance of the T-schema. However, since truth-talk is just a type of pretence, the truth-theoretic fictionalist is not forced to commit herself to the acceptance of true contradictions, but rather she is simply committed to the acceptance that certain contradictions are true according to the relevant fiction of truth.

It appears now clear what is the answer to the question that gives rise to this section, namely, "Should dialetheists be alethic fictionalists?". The question will have a negative answer, that is, dialetheists should not be fictionalists about truth-talk, and the reason is that dialetheists have no advantage in assuming a fictionalist perspective of the discourse about truth. As a matter of fact, if we compare the reasons of an advocate of inconsistency theories for embracing a truth-theoretic fictionalist perspective, it becomes clear that all these reasons

⁴⁰ Where saying that something is analytic to the notion of truth means that using the truth predicate with its common meaning commits us to the acceptance of that thing. For instance, one principle that is analytic of the notion of truth is the T-schema, or, more precisely, its instances.

⁴¹ Burgess[2007], p.45.

cannot be shared by the dialetheist and, on the contrary, that they correspond exactly to the main idea of a dialethic theory.

There are certain fundamental differences between dialetheism and inconsistency theories of truth that come to help us in the comprehension of why the latter need to embrace alethic fictionalism and the former does not. As Beall and Priest remember us, Eklund claims that “a correct account of truth is the one that comes closest to making the MC [i.e. meaning-constitutive] principles true, that is, one that maximally preserves the MC principles about the notion.”⁴² This can be easily achieved if we assume that truth is consistent, that is, if we assume that the law of non-contradiction is a fundamental pillar of the concept of truth. However, neither Priest nor Beall have this necessity, that is, they don’t need to maintain the law of non-contradiction as fundamental. On the contrary, they, as dialetheists, claim that the law of non-contradiction is far from being central of the concept of truth. The reason is well known:

A paraconsistent logician, if reasoning about a consistent situation, can use classical logic. [...] Classical logic is just the special case of paraconsistent logic [...] restricted to consistent situations. Now, the only situations that make sense from a classical perspective are consistent ones. Hence, reasoning about any situation that makes sense from a classical perspective makes sense from a paraconsistent perspective also. It is just that from a paraconsistent perspective there will be other situations about which we can reason, which make no sense from a classical perspective. The effects of rejecting the LNC, then, are not at all as drastic as one might have been supposed. [Beall and Priest 2007, p.81]

Therefore, the acceptance of the law of non-contradiction is not mandatory for a dialetheist, contrariwise to the advocates of inconsistency, who categorically don’t admit the possibility of true contradictions. In other words, dialetheists are not forced to take the law of non-contradiction as a key pillar of their theory. However, as we stressed again and again in chapter 1, the acceptance of the law of non-contradiction and the claim that truth is inconsistent are not mutually exclusive from a dialethic point of view. We can accept both the law of non-contradiction and the truth of certain contradictions as well. In other words, we can accept the law of non-contradiction even if we admit, at the same time, that there are true counterexamples to it. After all, this fact is not surprising at all because we know that the dialetheist admits, by definition, the possibility of true contradictions and, so, admitting both a law and its negation is perfectly possible for her.

As we saw, the main reason for advocates of inconsistency to adopt alethic fictionalism is that the latter allows them to keep their intuitions about the intrinsic inconsistency of the language safe, and, at the same time, it doesn’t force them to relapse into a more radical position that admits the existence of true contradictions *tout court*, i.e. outside the fiction. However, it is pretty clear that dialetheists don’t have this necessity because dialetheism is, by definition, the theory that there are true contradictions *tout court*. For this reason, dialetheists don’t need to embrace fictionalism about truth-talk in order to cut

⁴² Beall and Priest[2007], p.80.

an inconvenient ontological commitment as a result of their intuitions about truth and its behaviour. In other words, if fictionalism about a region of discourse has the virtue of avoiding an inconvenient ontological commitment towards the objects to which the sentences of that region of discourse refer to, then fictionalism about truth-talk is of no help for a dialetheist, because dialetheists don't need to avoid the ontological commitment towards truth, even when it is the source of true contradictions. This does not mean that dialetheism and alethic fictionalism are incompatible theories. Nothing prevents a dialetheist from endorsing a fictionalist perspective of truth-talk, but the point is that this is not mandatory. Moreover, the ontological beliefs are substantial on this matter. Dialetheism, as is well known, has a very rich ontology by definition, an ontology that includes also certain true contradictions. On the contrary, alethic fictionalism is an instrument to cut off the ontological apparatus of a theory. Therefore, dialetheism and truth-theoretical fictionalism are compatible, but there is no need for the dialetheist to make this marriage happen.

More on this point, both deflationary dialetheists, such as Beall, and Dummettian dialetheists, such as Priest, don't face the problem of the ontological commitment. If the dialetheist is deflationist, then – as we saw in section 4.2.1 – true contradictions arise only at the semantic level as inevitable consequences, i.e. spandrels, of the introduction in the language of the truth predicate that merely has the fundamental role of device of semantic descent. If, on the contrary, the dialetheist is not deflationist – and has, for instance, a Dummettian view of truth, according to which truth is the *telos* of assertion, and we know that this view requires a more robust nature of the property of truth than the one offered by the deflationist –, then dialetheias can appear also in the true-free language. This means that dialetheias appear independently from the introduction in the language of the truth predicate. In both cases, the dialetheists don't need to embrace a fictionalist view of truth-talk because, on the one hand, the thin nature of the truth predicate according to the deflationist, doesn't require the further assumption that truth doesn't exist *tout court*; and, on the other hand, the more robust nature of truth postulated by the Dummettian entails a more robust kind of true contradictions as well, but this is yet accounted by the theory and, for this reason, does not require a fictionalist perspective of truth-talk.

5.2.4 Conclusion

The starting point of the analysis was the following question: should dialetheists be fictionalists about truth-talk? Through the comparison between dialetheism and inconsistency theories of truth we determine a negative answer to the question: no, the dialetheist should not be fictionalist about truth-talk. The reason, however, is not, as one might think, that dialetheists cannot be alethic fictionalists because of some sort of incompatibility between the two accounts, but rather that there is no motivation for the dialetheist to assume a fictionalist perspective of truth talk. In other words, there is no advantage for dialetheists in endorsing alethic fictionalism. The reason why the advocates of inconsistency theory often adopt a fictionalist approach of the discourse about truth is to keep their intuitions about the truth predicate – that is, their

intuitions about the inconsistent character of truth talk – unchanged and, at the same time, to not being forced to accept the existence of truth contradictions. However, it is pretty clear that these reasons don't work for dialetheists: on the one hand, dialetheism, like inconsistency theories, admits the inconsistency of the truth predicate; but, on the other hand, it also admits the existence of true contradictions. Therefore, dialetheists don't need to embrace alethic fictionalism in order to avoid ontological commitment towards true contradictions because the existence of true contradictions is the main thesis of dialetheism itself.

5.3 Should Alethic Fictionalists Really Be Dialetheists?

If in the previous section I have reasoned starting from a dialethic point of view, now I'm going to invert the perspective in order to focus on fictionalism. In other words, if in the previous section I wondered whether the dialetheists should be fictionalists about truth-talk, or not, now I'm going to wonder the complementary question: should alethic fictionalists really be dialetheists as Beall claims?

The analysis so far has led to the conclusion that dialetheists have no advantages in the assumption of a fictionalist perspective of truth-talk. Here, in this section, I'm going to focus on alethic fictionalism in order to try to figure out whether the fictionalists about truth-talk find some advantages in assuming a dialethic theory, or not.

5.3.1 Background

The starting point of this analysis will be, as we have already said, the work of Beall made in the paper *True and False – As If*,⁴³ in which the philosopher develops his own specific account of truth, named *constructive methodological deflationism* (CMD). According to Beall, the truth predicate is merely a device of semantic descent, it doesn't correspond to any substantive property of truth (deflationism) and, more in details, truth is nothing but a construction that someone before us has introduced into the language in order to provide the language with an expressive device allowing us in the formulation of certain generalizations. This means that when we use the truth predicate within our language, we are “intentionally engaged in *as if* thinking,”⁴⁴ that is, we pretend truth to be a substantive property in order to say things that we cannot express otherwise without truth and that have nothing to do with truth itself (fictionalism). In other words, according to Beall, when we engage in truth-talk, we are, in fact, engaging in pretence, even if we are not aware. After all, the story of truth and falsity behaves in the same way as the story of Aiehtela and Aiehtelanu. Both stories are constructions, i.e. fictions, and the only difference between them is that in the case of truth we are not aware of

⁴³ Beall[2004].

⁴⁴ Ibid., p.206 (emphasis in original).

our *as if* attitude towards the sentences formulated within that area of discourse, whereas in the case of Aiehtela and Aiehtelanu we are always aware of this peculiar feature of the story. Moreover, the underlying logic governing the truth predicate, according to Beall, must be paraconsistent and the theory connected must be dialethic. This means, thus, that, according to Beall, an advocate of fictionalism of truth-talk has, in fact, certain advantages in the assumption of a dialethic theory that admits the existence of true contradictions and that makes it preferable over the other for her. The main benefit of dialetheism for Beall is that it allows fictionalism about truth-talk to face and solve the problem of the liar sentence.

However, it seems that the union between fictionalist deflationism and dialetheism is advantageous to both sides. In the first place, that truth is nothing but a fiction, i.e. a construction, is what makes unproblematic the admission of its inconsistency. According to Beall, it is pretty natural to think that a fiction is inconsistent, that is, it is natural to think that inside fiction things might go in a different manner than they go outside it, and inconsistency can really be one of those things. In fact, as Beall puts it:

Most constructions of the human mind tend towards inconsistency, especially if they aim towards completeness. [Beall 2004, p.213]

Hence, it seems that alethic fictionalism totally respects the dialethic intuitions that there are true contradictions. In fact, if it is natural to think of an inconsistent fiction, then it follows that assuming a fictionalist perspective of truth-talk is beneficial for the dialetheist that find a further motivation in it.⁴⁵

Moreover, from Beall's point of view, true contradictions are nothing but spandrels obtained as a result of the introduction into the language of this device of semantic descent that is truth and that is nothing but a mere construction useful only to accomplish the role it was introduced to. No wonder, then, that a construction has spandrels as a consequence. In addition, according to Beall, admitting the existence of true contradictions allows enjoying "a *complete* story of truth,"⁴⁶ that is, allows having truth and its principles without restriction. In particular, the admission of the existence of true contradictions enables us to have the full T-schema without being forced to restrict it in a way or another in order to preserve consistency. This means that if alethic fictionalists assume a dialethic theory of truth, they can solve the problem of the liar paradox in a very simple way without being forced to reject or restrict some very intuitive rules or principles that appear to govern the functioning of the truth predicate. This is possible exactly because of the admission of true contradictions. After all, if truth is a pretense, i.e. a construction, why do we want that fiction to be consistent? This is the main idea behind constructive methodological deflationism.

⁴⁵ Actually, in the previous section we saw that this is not the case. As a matter of fact, the dialetheist has no problem to admit that reality is inconsistent as well; hence she has no need to rely on alethic fictionalism in order to legitimate her intuitions. Beall's point on dialetheism still stand, that is, the idea that if the dialetheist assume a fictionalist perspective of truth-talk, he could do it but it is important to highlights that this is not a necessity for her, but an arbitrary choice.

⁴⁶ Ibid., p.208 (emphasis in original).

The question that guides this section, then, is the following: is Beall right when he says that a fictionalist approach of truth-talk finds great advantages in a paraconsistent logic and in a dialethic theory of truth, according to which there are true contradictions, or, on the contrary, this is not the case since if there are such advantages, they are not so decisive on the matter because other approaches have the same costs-benefits result? In other words, the question now is: should alethic fictionalists really be dialetheists?

5.3.2 Alethic Fictionalism is an Agnostic View

The purpose of this section is to figure out whether a fictionalist perspective of alethic discourse really must embrace a dialethic theory in order to solve the problem due to the semantic paradoxes, or not. Perhaps there are other alternatives to dialetheism for the alethic fictionalist such that dialetheism still remain a possibility for her, but not a necessity as someone argued.⁴⁷ In the words of Woodbridge:

In the instances of truth-talk, uses of expressions like “is true” and “is false” appear to attribute properties – truth and falsity – to objects that the term expressions supposedly denote. These appearances are just part of a pretense [...] There are no such properties as truth and falsity, and the expressions “is true” and “is false” do not even really play the linguistic roles they appear to play. We talk *as if* there are properties of truth and falsity in order to make certain serious assertions (not about truth) indirectly. [Woodbridge 2005, p.135 (emphasis in original)]

Starting from Woodbridge’s idea – i.e., the idea that we only pretend that expressions involving the truth and falsity predicates really denote the correspondent properties, but rather they are mere constructions useful only for the expression of things about the concrete world, that has nothing to do with truth and falsity –, may we say that Beall and Woodbridge are right when they assert dialetheism to be the only possibility available for the alethic fictionalist in order to solve the problem due to the liar paradox?

Beall and Woodbridge’s idea is that, in order to keep the full validity of the instances of the T-schema – that, remember, is a necessary requirement for a deflationary theory of truth –, we necessarily have to accept the inconsistency of the language, that is, we must accept that the truth predicate is governed by inconsistent principles. As Woodbridge puts the issue:

Taking the instances of (ES) [i.e. equivalence schema] as fundamental allows deflationary views to accept the *prima facie* paradoxical sentences as genuinely paradoxical and to pursue a strategy of diagnosing and containing truth-talk’s inconsistency, rather than one attempting to eliminate it [Woodbridge 2005, p. 138 (emphasis in original)]

According to the two philosophers, this means nothing but accepting our language to be inconsistent and the best way to do so is by relying on a dialethic theory of truth that admits the existence of true contradictions. After

⁴⁷ Beall[2004] and Woodbridge[2005].

all, Beall himself maintains that this result – i.e. the appeal to a dialethic theory by an alethic fictionalist – is not surprising for the following reason:

After all, if truth is a mere (human) construction, introduced to play a given expressive role, then it is not surprising – indeed, it is likely – that the construction should turn out to be inconsistent. [Beall 2004, p.208]

The idea, thus, is that alethic fictionalism shows in all its evidence that truth is ruled by inconsistent principles and that this does not come as a surprise because we are still talking of a fiction and there's nothing wrong or strange in an inconsistent fiction. On the contrary, it seems perfectly legit and even natural for a fiction not to be consistent.

These are the reasons provided by Beall and Woodbridge in favour of the union between alethic fictionalism and dialetheism. It is in the nature of truth and, in particular, in the nature of truth as a construction, i.e. as a fiction, to manifest this inconsistent character that, for this reason, must be completely accepted by relying on a theory that is able to keep it unchanged, that is, on a theory that is able to accept and account for the existence of true contradictions, i.e. dialetheism.

The idea I wish to support here is that it's not true that alethic fictionalism, as Beall and Woodbridge presented it – that is, as the view that truth involves elements of pretense, and not as the view that the claim that something is true means that we are pretending it is true –⁴⁸ should adopt a dialethic theory. Therefore, I wish to negatively answer the main question of the section. The reason, here as in the previous section, is not that dialetheism does not fit well with truth-theoretic fictionalism,⁴⁹ but rather that there are other possibilities available to the alethic fictionalist in order to provide an answer to the liar paradox, and these possibilities don't require a dialethic semantics. In other words, the idea is that the fictionalist does not need dialetheism in order to provide an adequate solution to the liar, contrariwise to what Beall and Woodbridge maintain in their work.

The first criticism to Beall's idea that is worth noting is due to David Liggins.⁵⁰ According to Beall, dialetheism is a perfectly natural choice to make for a deflationary fictionalist theory of truth such as his. In other words, according to Beall, in the very moment the main inconsistent character of truth arises, then there is no reason to look for a way to avoid this inconsistency. Within a fiction is perfectly acceptable that things behave in a contradictory way and, Beall adds, the opposite would be very odd. According to Beall, in fact, it would be odd if within a construction, i.e. a fiction, everything were consistent. Hence, in Beall's view the inconsistency of the truth predicate is perfectly natural in his theory. If inconsistency does not represent a problem

⁴⁸ To pretend something to be true would be, in turn, a fiction, because it would involve a further level of fiction in order to be evaluated. The idea, instead, is that "truth-talk functions in virtue of pretense, but speakers use it to say (indirectly) how things are, not just how they pretend things are." [Woodbridge 2005, p.135] Therefore we speak *as if* things has or lack properties called "truth" and "falsity", but we are actually making indirect claims about things that have nothing to do with truth and falsity.

⁴⁹ It is worth mentioning that Liggins[2014] has reservations about this regard. According to him, the realism of a dialethic theory conflicts with the spirit of a fictionalist perspective.

⁵⁰ Ibid.

within a fictionalist perspective of truth-talk, then it's clear that is not necessary to find a way to eliminate this inconsistency and, on this line, the most appropriate approach seems to be one that accepts true contradictions as natural, namely, dialetheism.

However, there are two things that are worth noting now. First, that truth turns out to be inconsistent within the fiction is perfectly natural in the same way as it would be natural that truth was inconsistent outside the fiction. However, this does not lead automatically to the acceptance of true contradictions. As we saw in the previous section, it is possible to assume a more moderate position about the acceptance of the inconsistency of the principles governing the truth predicate without postulating the existence of truth contradictions by virtue, for instance, of an inconsistency theory of truth, instead of the more radical dialetheism. Therefore, dialetheism is not the only possibility available to the alethic fictionalist who wishes to keep inconsistency rather than eliminate it. Secondly, as Liggins points out, "since Beall builds a dialethic logic into CMD [i.e. constructive methodological deflationism], how could CMD offer a non-circular motivation for dialetheism?"⁵¹ The two points are, thus, connected to one another. It seems, in fact, that adopting a dialethic logic is an arbitrary choice made by Beall and not a choice we are forced to make because of the inconsistent character of the fiction of truth, as Beall maintains. Beall claims that his deflationary and fictionalist account supports his own dialethic theory. It seems that dialetheism is the only choice for an advocate of his theory, but this is not the case. In other words, it seems that Beall adopted a dialethic logic from the beginning and that, only in a second moment, when he found out that truth within his fictionalist theory behaves in an inconsistent way, he maintained that this is a perfectly natural choice and, thus, a motivation in favour of a dialethic logic. However, he could maintain that only because he assumed a dialethic logic from the beginning and for this reason, as Liggins highlights, Beall's idea is based on a circular argument.

Moreover, Liggins notices that in Beall's argument there is an implicit and illicit step from what we can say within the fiction to what we say outside it. Beall's argument for the existence of true contradictions is the following:

- | | | |
|------|--|------------------------|
| (P1) | Either (L) is true or (L) is false. | [LEM] |
| (P2) | (L) is true iff (L) is false. | [instance of T-schema] |
| (C) | Therefore, (L) is true and (L) is false. | |

where (L) is the liar sentence.

The problem in this argument is that since truth is a fiction, for the argument to be valid it is necessary for both its two premises and the conclusion to have a prefix specifying the relevant fiction. For instance, the correct formulation of (P2) must be the following: "according to the story of truth, (L) is true iff (L) is false." Hence, all we can correctly infer is that according to the fiction of truth there are true contradictions, but this doesn't mean that there are true contradictions *tout court*, contrariwise to dialetheist's main claim. The reason is that we can't legitimately move from "according to the story f, p" to "p". Therefore, the fact that the fiction of truth is inconsistent

⁵¹ Ibid. p.569.

provides no motivation in favour of a dialethic theory of truth. As Liggins puts it:

Beall claims that the story (“construction”) of truth is likely to be inconsistent. [...] Perhaps so. But to infer that there is likely to be a true inconsistency is to perform an invalid exportation. The most we will be able to infer is that, *according to the story of truth*, $A \ \& \ \sim A$. [Liggins 2005, p.571 (emphasis in original)]

This does not mean that (P1) and (P2) are wrong, as a revolutionary fictionalist would claim, but rather that both are right and we are disposed to accept that they really govern the fiction of truth only if we interpret them in the right way, that is, in an hermeneutic fictionalist spirit.

In line with what we said, there is a further remark that is worth mentioning if we follow Liggins’ argument. As we stressed, the dialethic fictionalist does not assert the instances of the T-schema in the following formulation:

(T1) “A” is true iff A.

where A can be any sentence of the language. But, in order for the T-schema to be asserted and accepted by the alethic fictionalist, it needs to be anticipated by the right fictional prefix. Hence, an advocate of deflationary fictionalism asserts and accepts only the following weaker formulation of the T-schema:

(T1*) According to the story (pretence) of truth, “A” is true iff A.

The point, here, is that the truth-theoretic fictionalist rejects (T1) and, by so doing, she can classically solve the problem of the liar paradox without resorting to dialetheism. As a matter of fact, by rejecting to assert the unprefixated T-schema, the argument of the liar can be easily blocked at the second premise, (P2). As a result, the derivation of the contradiction becomes impossible. But if the liar can be solved in this classical way, then it very clearly appears that dialetheism becomes useless for an alethic fictionalist, because the main motivation for its assumption, in fact, fails.

Lastly, besides the classical solution and the one provided by inconsistency theories of truth, there is another possibility to solve the liar paradox available to the alethic fictionalist. The solution, suggested by Armour-Garb and Woodbridge,⁵² is a version of the meaningless strategy, according to which the liar lacks of content. In Armour-Garb and Woodbridge’s fictionalist account, the two philosophers define the fundamental concept of *M-conditions*, that is, what we need in order to grasp the “wordly-content aspect of sentences”.⁵³ In other words, the M-conditions are those conditions of a sentence specifying how the real world is according to that sentence. The account of make-believe developed by the two authors makes possible the indirect specification of M-conditions through the use of a fictional claim, instead of allowing the direct specification of the M-conditions through the use

⁵² Armour-Garb and Woodbridge[2015b].

⁵³ Armour-Garb and Woodbridge[2015a], p. 48.

of a claim not involving the pretense. Roughly speaking, Armour-Garb and Woodbridge are maintaining that the liar lacks of content because it specifies no M-conditions. More in details, to determine the M-conditions that a fictional sentence indirectly specifies, we need to look at the M-conditions the correspondent sentence not belonging to the fiction directly specifies. In the specific case of the liar, however, the correspondent sentence is the liar itself and, hence, the process will repeat endlessly. The result, then, is that the liar cannot determine any M-condition and, in fact, this is equivalent to what a meaningless strategy claims with respect to the liar.⁵⁴

In conclusion, the current analysis has proved that there are different ways to solve the liar paradox available to an advocate of a truth-theoretic fictionalist perspective. In fact, besides the dialethic theory, she can rely on an inconsistency theory of truth, on classical logic or on the meaningless strategy provided by Armour-Garb and Woodbridge. Hence, dialethicism is far from being necessary for an advocate of alethic fictionalism.

5.3.3 Conclusion

The question guiding this section was the following: should alethic fictionalists really be dialetheists? Beall affirmatively answers the question: dialethicism allows truth-theoretic fictionalism to solve the problem of the liar by keeping the intuitions about the inconsistency of the language and the behaviour of the truth predicate in it unchanged. According to Beall, truth is nothing but a useful construction helping us in the expression of things we couldn't express otherwise. This is why Beall defines his theory as a deflationary fictionalist theory: it is a fictionalist theory because to say that truth is a construction means that truth is a fiction or, in other words, that truth works through elements of fiction; and it is a deflationary theory because to say that there is nothing on the truth predicate but what allows us to use it as a device of semantic descent, i.e. the T-schema, just means endorsing the deflationary main thesis. The inconsistent character of truth appears immediately when we deal with the liar and, for this reason, the dialetheist offers to the alethic fictionalist an appropriate account for the resolution of the liar and for keeping the intuitions about the inconsistency of the truth predicate in the language.

My aim in this section was to motivate the idea according to which the beginning question should have a negative answer, namely, that alethic

⁵⁴ However, every meaningless strategy faces the same objection: if the liar really were meaningless, then we would not understand it, but, in fact, we understand it, hence the liar cannot be meaningless. Armour-Garb and Woodbridge provides an answer to this objection. In particular, the two philosophers maintain that there are two different notions of "understanding". The first, named "understanding₁", requires the specification of M-conditions, whereas the other, named "understanding₂", corresponds to the comprehension we have of the liar sentence. The latter is a metalinguistic kind of comprehension that presupposes us to comprehend what a sentence seems to be, namely its form, its subject, its predicate, and so on. More in details, with this kind of comprehension we, in fact, know "how such a sentence could be used to make a meaningful assertion." [Armour-Garb and Woodbridge 2015b, p.346]. Armour-Garb and Woodbridge's idea is that if we know all those things about a sentence – its form, the meaning of each word and the way we can use it to express a meaningful assertion –, then we will understand₂ the sentence. Since, however, we don't know the M-conditions of the sentence, then we cannot say that we understand₁ the sentence. Therefore, on this view, we understand₂ but not understand₁ the liar sentence.

fictionalists shouldn't be dialetheists, contrariwise to what Beall claims. The reason is not that dialetheism does not fit well with a fictionalist account of truth-talk,⁵⁵ but rather that there are other theories available to the fictionalist in order to solve the liar.

As a matter of fact, as the analysis pointed out there are at least three others theories providing a solution to the liar and fitting very well with alethic fictionalism. In the first place, the inconsistency theories of truth, because the liar shows that inside the fiction the truth predicate is ruled by inconsistent principles, but this means only that according to the story of truth there are true contradictions, and since we cannot automatically move from a sentence formulated inside the fiction, i.e. with fictional prefix, to the correspondent sentence without fictional prefix, i.e. outside the relevant fiction, then from the alethic fictionalist account doesn't follow that there are true contradictions *tout court*, that is exactly the thesis defended by dialetheists. On the contrary, if we are truth-theoretical fictionalist, we can assume a more moderate position, namely, an inconsistency theory of truth. In the second place, classical theories of truth, because the instance of the T-schema used to derive the contradictory conclusion in the liar reasoning can be interpreted in two ways, either it is an instance of (T1*) – i.e. an instance of the T-schema formulated inside the fiction – and in this case we are rebounding on the previous alternative of inconsistency theories and we are not talking of true contradictions anymore, or it is an instance of the T-schema formulated outside the fiction, i.e. (T1), but, then, that instance is not accepted and even asserted by the fictionalist of truth-talk and, in virtue of this rejection she can block the derivation of the liar argument in a classical way. Lastly, alethic fictionalist can adopt a meaningless strategy if she maintains that the liar lack of content, as Armour-Garb and Woodbridge wishes to say, because it doesn't specify the necessary conditions to determine it, i.e. M-conditions.

In conclusion, there's nothing wrong in the claim that dialetheism is a good theory for a fictionalist account of truth-talk. However, if one wishes to take the point one step forward and claims that the alethic fictionalist need to assume a dialethic theory of truth, then she will be drawing a misleading conclusion. In other words, if the question is "Can alethic fictionalists be dialetheists?", then the answer is without doubts affirmative, since there is nothing in the fictionalist perspective that conflicts with dialetheism. However, here, the question is another, namely, "Should alethic fictionalists really be dialetheists?", and this, as we saw, changes the answer to negative. No, the fictionalist about truth-talk should not be dialetheist because she doesn't have the necessity to endorse such a theory. As a matter of fact, she has equal good reasons to endorse one of the following theories: inconsistency theories of truth, classical theories, a meaningless strategy or a dialethic theory.

Hence, to decide which is the best theory for a fictionalist account of truth-talk is to decide which is the best theory of truth *tout court* – at least among the four I just mentioned – independently from the theory of the discourse about truth we endorse.

⁵⁵ As we saw in footnote 49, Liggins suggests there is a tension between the realism of a dialethic theory and the spirit of a fictionalist view of truth-talk.

5.4 A New Perspective About The Liar

5.4.1 Introducing Yablo's Account

So far I have taken into account Beall's attempt to keep all the three theories we analysed – i.e. deflationism, fictionalism and dialetheism – together and we concluded that his attempt doesn't work for several reasons, *in primis* the fact that dialetheism is far from being necessary for a fictionalist theory of truth-talk.

There is an argument we mentioned at the end of chapter 3 and that we have never discussed yet, namely, Armour-Garb and Woodbridge's argument according to which the deflationist should be fictionalist about truth-talk. The argument makes an analogy with Yablo's argument about mathematical discourse. Armour-Garb and Woodbridge maintain that if the reasoning made by Yablo on fictionalism about mathematical discourse works – namely, that nominalism about mathematical objects should assume a fictionalist perspective of mathematical-talk –, then the similar reasoning they make for alethic fictionalism must work as well – namely, that a deflationary theory of truth should assume a fictionalist perspective of truth-talk. The argument hinges on the idea that both numbers and truth work as representational aids and that working as representational aids means working through some elements of fiction. Hence, the point is that, as mathematical nominalism works through elements of fiction and for this reason adopts a fictionalist perspective of mathematical-talk, in the same way deflationism works through elements of fiction as well and, hence, it should assume a fictionalist perspective of truth-talk.

My purpose in this section is to take the analogy with Yablo's theory one step forward and try to outline a proposal of how it is possible to apply his recently account of subject matter to the problem of the liar and how it can be translated into the fictionalist vocabulary.

In the book *Aboutness* published by Yablo in 2014, the philosopher develops a new strategy in the understanding of the meaning of a sentence in a theory involving certain particular notions, such as those of "subject matter" and "partial truth". After having exposed its main features, I'm going to apply the theory to the problem of the liar paradox, in order to see if it can help us in its resolution. In order to do so, I wish to show that the main concepts of the new account of subject matter can be translated into the main concepts of fictionalism and, hence, that this last work by Yablo is not so far away from his previous ones, but rather it is only their more sophisticated development.

5.4.2 Semantic Excuses

Yablo's account of aboutness can be understood as a new way of thinking about meaning, a new way that, perhaps, has the advantage to having found an agreement even between the most different metaphysical views, such as nominalism and Platonism in philosophy of mathematics.⁵⁶

⁵⁶ I'll come back to this point at the end of the chapter.

The real focus of Yablo's account is the relation of aboutness – as the name of the book suggests –, but in order to understand this notion we must go first through certain intermediate steps in order to define certain fundamental and preliminarily concepts useful for the comprehension of the whole account. We should, thus, define first the notion of “partial truth” and then the notion of “subject matter” and finally we can explain what the relation of aboutness really is.

Actually, the proper starting point in order to explain Yablo's idea and its functioning is a different concept, namely, semantic excuses. Semantic excuses are excuses we provide in order to assert things that are or might be false.⁵⁷

Let's begin by imagining a hypothetical situation where Sally is discussing her own dissertation on the topic of physical objects and their identity over time. The argument provided by Sally presented a problem. In fact, in order to talk about persistence of physical objects over time, one inevitably has to speak about gaining and losing of properties. However, Sally didn't want to take a position on the metaphysics of properties, that is, she didn't want to talk about the existence or non-existence of properties. When Sally, for instance, said something like: “the tomato loses the property of being green and gains the property of being red”, she was not meant to express any sort of ontological commitment towards the properties of redness or greenness. The problem is that the utterance made by Sally is false if properties didn't exist; hence, it seems natural to think that with an assertion of that kind she was committing herself to the existence of properties. But she was absolutely unbothered by the possibility that properties didn't exist. How is it possible? Her semantic excuse to explain the situation was that the topic of her statement and, more in general, of its entire dissertation were not properties, but material objects and their persistence through change.

The question now is: how can this be an acceptable excuse for asserting a falsehood? Or, in other words, can this – i.e. to say that Sally didn't want to speak about properties because the real topic of her dissertation were not properties but physical objects and their persistence over time – be a convincing semantic excuse for asserting a falsehood and, in particular, for asserting sentences referring to properties and that, for this reason, are false?

After all, it is enough for Sally that her utterance of “the tomato loses the property of being green and gains the property of being red” was true in what it says about the tomato and its persistence over time. She was not interested in the issue of the existence of properties and, so, she was not interested in making the sentence true *tout court*. This fact suggests that sometimes, in certain specific contexts, asserting a falsehood can be excused because it is sufficient for the assertion of the sentence that the sentence is only partially true, that is, true for what it says about its real topic. Or, in other words, the point is:

It might be that certain truths are not accessible except as scattered parts of larger falsehoods. [Yablo 2014, p.77]

This is Yablo's main idea in the development of his account.

This allows us to provide an answer also to the above questions. To say that the topic of the statement made by Sally are physical objects and their

⁵⁷ Yablo[2014], p.77.

persistence over time is, according to Yablo, an acceptable semantic excuse for asserting a falsehood. The reason is that certain truths are not expressible except as parts of larger falsehoods. If we want to express a truth and the only way we have to do so is by asserting a falsehood, then, according to Yablo, we are legitimate to express it in that very way. The idea is that in all these cases we can individuate the part of the content of a false sentence *A* that is true about a specific subject matter, but we can't have a sentence expressing it. In other words, we can individuate the true proposition that is important to us but we have no way to express it directly. We can only indirectly express that true proposition as part of a larger falsehood. In this way, then, asserting a falsehood can be sometimes a right price to pay if it is the only way we have to express a truth.

This intuitively is Yablo's idea. The great advantage of this account is that it applies to several different domains. Let's now display few other examples from different regions of discourse in order to clarify the functioning of this particular mechanism that is partial truth.

The first emblematic example is about the so-called loose talk. It's interesting to see the functioning of Yablo's account with regard to this region of discourse because it's an area commonly used by almost everyone that does not require specialist knowledge in order to be employed. If someone asked about Yablo's height, we could correctly answer that Yablo is 5 feet and 9 inches tall. But this answer is not true because Yablo is closer to 5 feet and 8 $\frac{3}{4}$ inches than to 5 feet and 9 inches. However, it's clear that he can't be both less than 5 feet and 9 inches tall and 5 feet and 9 inches tall. How can we explain the fact that the sentence "Yablo is 5 feet and 9 inches tall" still sound right? The reason is that the sentence is true about height in inches, that is the most common unit of measurement for height in the U.S. Hence, the sentence "Yablo is 5 feet and 9 inches tall" is false because Yablo is less than 5 feet and 9 inches tall, but it is partially true because it's true about height in inches – and not about height in fractions of inches – that is the relevant topic of the sentence.

The same goes for another sentence from loose talk: "France is hexagonal". The sentence is false if we considered the real shape of France, but we can still say that it is true, i.e. it is partially true, if we consider the relevant subject matter of the sentence that is the approximate shape of France.

The third example I'm going to mention is about a different region of discourse, namely, applied mathematics. Let's consider the well-known sentence: "The number of Martian moons is two." If we assume nominalist perspective of mathematical objects, then this sentence is false because on this perspective there are no such things as numbers. However, if we analyse the sentence we can figure out that with an utterance of that sentence we are not talking about numbers, but rather we are talking about Mars and its moons. Hence, the sentence is false, but it is partially true, that is, it is true about Mars and its moons and, more in general, it is true about the concrete world.⁵⁸

⁵⁸ On this regard it is worth noting that a very interesting feature of Yablo's account, that I already mentioned before and that appears very clearly now, is that Yablo's account remains neutral with respect to the metaphysical dilemma between the Platonist and nominalist views in philosophy of mathematics. As a matter of fact, the sentence "the number of Martian moons is two" is true both platonistically and nominalistically if we consider the relevant subject matter of

Finally, the last example that it is worth mentioning in order to show that certain truths can be conveyed only as partial truths of a larger falsehood is borrowed from the discourse about non-existence objects. Let's consider the famous sentence "Pegasus doesn't exist". Pegasus is a fictional character and, thus, it doesn't exist, and if we are not Meinongianist, Pegasus doesn't subsist as well.⁵⁹ For these reasons the sentence turns out to be false, because it involves terms without reference, i.e. "Pegasus". However, our intuitions about the sentence bring us in the opposite direction, that is, we intuitively want to say that the sentence is true because, in fact, there is no object such as Pegasus. How can we explain our intuitions together with the fact that the sentence is false? And, how can we explain the difference between the following sentences: "Pegasus does not exist" and "Pegasus exists"? Both sentences are false, because both sentences involves a non-referring term, but there is a sense according to which we wish to say that the former is true and the latter is simply false. How can we explain this difference? In other words, in order to prevent the two sentences from generating a contradiction, a specific analysis of the two sentences must be provided.

Let's consider a sentence similar to the previous: "Pegasus is not in the room". This sentence, as the previous ones, is false for the same reason, i.e. "Pegasus" refers to nothing. However, Yablo tells us that the sentence is false but also partially true because it is true with regard to how things are about the room and its content. We can make the same reasoning for the sentence "Pegasus does not exist". The latter can be paraphrased by the following sentence: "Pegasus is not one of us", where with the word "us" we mean all existent objects. Therefore, by applying Yablo's account, this sentence is false but partially true, that is, it's true about the relevant subject matter that, in this specific case, is "us", i.e. all existent objects, because it is true that Pegasus is not among existent objects. We can now explain the difference between the two sentences "Pegasus does not exist" and "Pegasus exists". The latter, similarly to the former, can be paraphrased by the sentence "Pegasus is one of us" and this is false and it is not even partially true because it's false also about the relevant subject matter, that is, it's false also about us, i.e. the existent things. In other words, among existence things we cannot find anything that is Pegasus and hence, "Pegasus does not exist" is true about that topic and "Pegasus exists" is false *tout court*, also about that same topic.

All these examples are useful to show the behaviour of the intuitive concept of partial truth, namely, of sentences that are false but have a true part. Those truths can be expressed only as a part of a larger falsehood. The examples of application are not limited to those, but rather they cover several areas of discourse, among the others it is worth mentioning pure mathematics, intentional identity, narrow content and the laws. In all those cases, we can find

the sentence, that is the concrete world. This means that both a nominalist and a Platonist could endorse such an account because it is neutral about the intuitions of either.

⁵⁹ The Austrian philosopher Alexius Meinong maintains that non-existent objects must have some sort of being since it seems possible for us referring to them and saying true things about them. For instance, it's possible to truly say that unicorns have horns, even if they are non-existent objects. For this reason, Meinong concludes that they must have some sort of being, different from that of existent objects but still a being. To maintain the differences between the two categories of objects, Meinong says that the property belonging to non-existent objects is the subsistence, whereas the property of existent objects is existence. [Meinong 1960]

out the true proposition we wish to express but we don't have a sentence to actually express it. The only way to convey the truth is to assert a false statement involving a true part. As Yablo puts it:

If access were limited in this way, then dallying with the larger falsehoods could be a good on balance a good policy. [Yablo 2014, p.77]

5.4.3 Partial Truth

After having listed a series of sentences, from different areas of discourse, that can be identified as semantic excuses – i.e. false sentences that we utter in order to express a truth included in them that cannot be expressed otherwise – we can, now, see the technical details of Yablo's account.

The first point to take into account is the definition of partial truth. Yablo takes as definition of partially true hypothesis the following:

(PT) A hypothesis is partly true iff it has parts that are wholly true.⁶⁰

The choice to use the word “hypothesis” is not accidental, but it is intentionally used to indicate in an ambiguous way both sentences and their propositional content.

If partial truth is defined as being a wholly true part of a larger hypothesis, then the second point to take into account will be to clarify what Yablo means when he speaks of an hypothesis that is part of another, that is, we need to define the relation of inclusion, or parthood, among hypotheses.

The most naïve definition of part is the following:

(PN) A hypothesis B is part of another hypothesis A iff B is implied by A.

Hence, A includes B iff A implies B, where both A and B are hypotheses.

A problem immediately arises. As a matter of fact, it seems that the relation of inclusion cannot be reduced to mere implication. Let's consider, for instance, the conjunction and the disjunction. In the case of conjunction we have a case of implication that is also a case of inclusion. The conjunction, hence, is a paradigm of inclusion: we know, in fact, that the conjunction implies each of its conjuncts ($\mathbf{p \wedge q \rightarrow p}$) and that the conjuncts are part of the conjunction ($\mathbf{p \leq p \wedge q}$).⁶¹ However, things are different with regard to disjunction. In this case we have a case of implication that is not also a case of inclusion. For this reason, we can define the disjunction as a paradigm of non-inclusion: we know that a disjunct implies the disjunction ($\mathbf{p \rightarrow p \vee q}$), but it is pretty obvious that the disjunction is not included in the disjunct ($\mathbf{p \vee q \not\leq p}$). Otherwise, we would have the following awkward consequence: the false sentence “ $2+2=5$ ” implies the true disjunction “ $2+2=5 \vee 2+2=4$ ”, and, if all implications were paradigm of inclusion, then “ $2+2=5 \vee 2+2=4$ ” would be

⁶⁰ Yablo[2014], p.11.

⁶¹ I'm using the same notation used by Yablo in the book, where \leq is the symbol for inclusion. What is on the left is included in what is on the right.

part of “2+2=5” and the latter would be partially true in virtue of the truth of the former, but this clearly is an absurdity. Hence, both conjunction and disjunction are cases of implication, but only conjunction is also a paradigm of inclusion. This means that all inclusions are also implications, but not all implications are also inclusions. Therefore, the relation of inclusion is stronger than implication. This result, however, is not predicted by the naïve definition of being part. Hence, (PN) cannot be an adequate definition of the relation of parthood.

What is missing in the naïve definition of parthood? According to Yablo, the problem with (PN) is that it does not take into account the explanatory role of the parthood relation, which requires more than mere implication in order to be performed.⁶² Good evidence of this can be found in several examples. For instance, let’s consider an example from *saying*: if I say “snow is white and expensive”, I’m also saying that snow is white; on the other hand, if I say “snow is white”, I’m not also saying that snow is white or expensive. However, they are both cases of implication. The second example is from *musts*: ordering someone to eat pork chops is ordering her to eat pork; on the other hand, ordering someone to eat pork is not ordering her to eat pork or human flesh. However, they are both cases of implication. The third example comes from *knowledge*: looking at a ripe tomato tells me that it is red, but not, that the tomato doesn’t misleadingly appear to be red. However, the latter is a consequence of the former. How could it be? Finally, if we consider *partial truth*, we will see that the same goes for it as well: the sentence “snow is white and expensive” is made partly true by the fact that snow is white. On the other hand, “snow is expensive” is not made partly true by the fact that snow is white or expensive. All these examples show that parthood is a relation stronger than implication. The difference is that true parts give partial true to their whole, whereas other kinds of implication don’t.

The question now is the following: which is the lacking ingredient that, together with implication, gives us the relation of parthood? We are looking for the secret ingredient that allows us to complete the following equation:

$$\text{Parthood} = \text{implication} + X$$

What’s the X? Yablo’s proposal is to substitute the X with the subject matter of a hypothesis. Therefore, in order to speak of the part of a sentence we must find an implication where the subject matter of the consequent is included in the subject matter of the antecedent. In Yablo’s words, the idea is the following:

The proposal is that for B to be part of A involves, in addition to A implying B, that B’s subject matter be part of A’s subject matter. [Yablo 2014, p.14]

Hence, the reason why certain implications of a hypothesis A are not included in A itself, like in the case of disjunction, is that their subject matter is completely foreign to that of A.

With this in mind, the definition of parthood becomes the following:

⁶² Yablo[2014], p.12.

- (PS) B is part of A iff the inference from A to B is such that both:
- (i) A implies B; and
 - (ii) A's subject matter includes that of B.

Therefore, the relation of parthood must be truth-preserving, because it is a case of implication, and it must also aboutness-preserving, because it must preserve the subject matter.

5.4.4 Subject Matter

In the previous section we defined partial truth in Yablo's account and, in order to provide such a definition, we incurred in the relation of parthood between hypotheses. The relation of parthood, according to Yablo, is not limited merely to implication, but it is a stronger relation involving the concept of subject matter of a sentence. In this section, I'm going to take the analysis one step forward and focus on this new concept of subject matter.

The first thing that comes into our mind when we think about the subject matter of a sentence is its truth conditions and we might wrongly think that the two concepts are equivalent. However, this is not the case. In fact, Yablo is very clear on this point and he stresses that truth conditions of a sentence only underdetermine its subject matter. More in details, Yablo maintains that a sentence's subject matter is a potentially independent factor because it can vary even if its truth-conditions remain the same.⁶³

To emphasise the difference between the two concepts, we can consider a very famous example, i.e. Hempel's raven paradox. Let's consider the following two sentences: "all ravens are black" and "all non-black things are non-ravens". They have the same truth-conditions and, in fact, they are true under the same circumstances. For this reason, one could think that the data confirming one, confirm also the other. However, this is not the case. As a matter of fact, the existence of a black crow seems to be more relevant for the confirmation of "all ravens are black" rather than the other sentence, whereas the existence of a non-black non-raven seems to be more relevant for the confirmation of "all non-black things are non-ravens". More in general, let's consider the following two sentences we assume to be true: "All Fs are Gs" and "All Gs are Fs". They are true under the same circumstances, hence they have the same truth-conditions, but they are about different things. The former is about the world's Fs and which among them are G; the latter is about the world's Gs and which among them are F. The answer provided by Yablo to this problem is that in both cases the truth conditions of the sentences are the same because the sentences are true under the same circumstances; what changes is the subject matter of the sentences. The two sentences are confirmed by different data because their subject matters are different. In fact, "all ravens are black" is about ravens and their colour and it's not about the colour of desks or whether non-black things are ravens or not. On the other hand, the sentence "all non-black things are non-ravens" is about red desks and all other non-black things that are not ravens, rather than the colour of ravens. Similarly for the Fs and

⁶³ Yablo[2014], p.19.

Gs: “all Gs are Fs” is about the Gs, whereas “all Fs are Gs” is about the Fs, despite their truth conditions being the same.

We can, thus, conclude that the subject matter of a sentence does not coincide with its truth conditions. However, we didn’t know yet what a subject matter is.

The most intuitive way to represent a subject matter is like a topic. What is common to all topics is that we can find a question for each of them. Hence, we can represent the subject matter also like an indirect question. For instance, if we consider the subject matter **what you did last summer**,⁶⁴ we can find the corresponding direct question that is *What did you do last summer?*

It’s worth noting that representing subject matters like questions is very useful because helps us to grasp also the intuitive idea of inclusion between subject matters. For instance, in order to figure out whether the subject matter **what you did last July** is included in the subject matter **what you did last summer**, or not, we have to consider the respective direct questions. The direct question of the former is *What did you do last July?* and the direct question of the latter is *What did you do last summer?* Providing a full answer to the latter entails providing an answer also to the former; hence, the latter subject matter includes the former.

If the intuitive idea of a subject matter is of picturing it as a question, the step forward is to provide a more rigorous definition of the concept. To provide such a definition Yablo uses the idea of subject matter developed by David Lewis.⁶⁵

First, according to Lewis, there are two different types of subject matter: the subject matter based on parts and the subject matter not based on parts. The subject matter **m** is based on parts if “for worlds to be alike with respect to **m** [i.e. a subject matter] is for corresponding parts of those worlds to be intrinsically indiscernible.”⁶⁶ The subject matter **nineteenth century** is based on parts because worlds are similar when **nineteenth century** is involved if and only if the nineteenth century of a world is an intrinsic duplicate of the nineteenth century of the other.⁶⁷

Among the subject matters not based on parts there are the following: **the number of stars** and **the observables**. With regard to **the number of stars**, there is no stars-counter such that two worlds agree about the number of stars if and only if the stars-counter of a world is an intrinsic duplicate of the stars-counter of the other. Similarly, things go for the subject matter **the observables**. Worlds are equivalent about the observables – namely, what can be seen, heard, etc. – if and only if their observables are intrinsically identical. The problem is that even if apparently two worlds appear perfectly identical

⁶⁴ I’m going to use bold font for subject matter.

⁶⁵ Lewis[1988].

⁶⁶ Yablo[2014], p.25.

⁶⁷ As we already said, the bold font indicates that we are speaking of subject matters. The difference between “the nineteenth century” and “**the nineteenth century**” is that with the former we are talking about the part of a world and its history, whereas with the latter we are talking about the way we can group worlds on the basis of how things stand about their nineteenth centuries. Another example can help to clarify this point: Albert was married to the queen and not to **the queen** because the queen is a person and **the queen** is a subject matter and one cannot marry a subject matter.

with regard to their observables, they could differ in their subatomic structure. This is why the subject matter **the observables** is not based on parts.

Is it possible to have a unique notion of subject matter including both the subject matters based on parts and those not based on parts? And how? According to Yablo, in order to have a unique notion of subject matter we must move from speaking of parts to speaking of partitions, because the latter is a more general notion than the former.

A partition, in fact, is a way we can group different worlds. Referring to the notion of partition we can finally provide a rigorous definition of subject matter:

(SM) Subject matter =_{df} a partition of the set of worlds.⁶⁸

Then, the identity relation between two subject matters is the following:

(IS) $\mathbf{m}_1 = \mathbf{m}_2$ iff worlds differing where the first is concerned differ also with respect to the second, and vice versa.

A subject matter is a partition of the set of worlds and a partition is a decomposition of some set into mutually disjoint subsets, called *cells*.⁶⁹ The **m**-cell of a world can be defined also as the proposition telling us how things are in that world with respect to **m**. Hence, a subject matter can be defined in different way: it is a partition of logical space and it is also a set of propositions expressing the various ways matters can stand where **m** is concerned.⁷⁰

Which is the proposition that specifies how matters stand in the world *w* where the subject matter **m** is concerned? This proposition is true in all and only the worlds in the same **m**-condition as *w*. Therefore, the proposition that specifies how matters stand in the world *w* where the subject matter **m** is concerned, is the set of worlds in the same **m**-condition as *w*. Let me clarify the issue with an example. To find the cell **the number of stars** of a world we must find a proposition specifying how things stand in that world when **the number of stars** is involved. Which is the proposition specifying how matters stand in the world *w* where the subject matter **the number of stars** is concerned? Let's suppose *w* has a billion stars. The cell **the number of stars** of *w* is the set of worlds with exactly as many stars as *w* and, hence, the cell **the number of stars** of *w* is the set of worlds with exactly a billion stars. Therefore, the worlds with a billion stars compose the proposition that there are a billion stars.

5.4.5 Aboutness and Sentential Subject Matter

In the previous section we defined the general notion of subject matter as a partition of logical space, that is, a partition of the set of worlds. In other

⁶⁸ Yablo[2014], p.27.

⁶⁹ Ibid., p.26, footnote 5.

⁷⁰ As a matter of fact, propositions are sets of worlds and the subject matter is a partition of the set of worlds, hence, the subject matter is a partition of propositions and, thus, the subject matter can be conceived as a set of propositions.

words, the subject matter is a set of propositions. Let's now see what it means for something to be true about a specific subject matter.

Previously I emphasised that Yablo's main idea is that a sentence can be false but partially true, namely, that can contain a part that is wholly true. But, what does it mean for A to be true about a specific subject matter **m**?

In the words of Yablo:

A is true about **m** iff one can make A true outright without changing how matters stand where **m** is concerned. [Yablo 2014, p.32]

In other words, if changing everything we can change about A, by keeping **m** unchanged, it is possible to make A wholly true, then A is true about **m**. More technically, this can be translated as follows:

(Im) A is true about **m** in the world w iff A is true simpliciter in a world **m**-equivalent to w.

This means that if we change world, by keeping unchanged how things stand about the subject matter **m**, and if we find a world where A is wholly true, then it follows that A is true about **m**. The reason is that A is wholly true in a world where **m** remains unchanged and, hence, in that world A is true also about **m**.⁷¹ For instance, if we consider the sentence "The number of Martian moons is two". The relevant subject matter for the evaluation of this sentence is **concreta**. Is this sentence true about **concreta**? We must find a world in the same cell as ours, that is, in which things are in the same way as in our world about **concreta**, and that make the sentence true simpliciter. A world identical to ours except that there are numbers is what we are looking for. If there is such a world, then "The number of Martian moons is two" will be true about **concreta** in our world.

Now we can finally speak about the relation of aboutness for Yablo. According to Yablo, *aboutness* is the relation between a sentence and its subject matter.

So far we saw what is a general subject matter and we defined it as a partition of logical space; moreover, we saw what means for something to be true about a subject matter; what is left is to define the notion of *sentential subject matter*.

The first point that is worth noting is that the subject matter of a sentence S can't be **whether S is true**. If to determine the subject matter of a sentence were sufficient to figure out if that sentence is true, then the relation of parthood of a sentence would be equivalent to implication. But we saw that the two relations must be kept separated because the former is a relation stronger than the latter, since parthood involves an explanatory role that implication doesn't perform.

⁷¹ This allows Yablo to define the phenomenon of what he calls *quasi-contradiction*. In fact, both a hypothesis and its negation can be true in the same world about the same subject matter. S is true about **m** in the world w if there is a world w' **m**-equivalent to w in which S is true simpliciter; and $\neg S$ is true about **m** in the world w if there is a world w'' **m**-equivalent to w in which $\neg S$ is true simpliciter. This is the phenomenon of quasi-contradiction. [Yablo 2014, p.33]

Let's try to clarify this point through an example. Have a look at the following sentence:

- (1) The U.S. president in 2001 is a senator's son.

Let's suppose, now, that in the world w' , the president is Dubya, son of senator George H.W. Bush and in the world w'' , the president is Al Gore, son of senator Albert Gore Sr. (1) turns out to be true in both cases, hence if the subject matter of the sentence (1) is **whether (1) is true**, then there won't be any difference between the two cases. However, we realize that, in fact, there is a substantial difference between them, because there is a change in the person that becomes the U.S. president. Therefore, what changes in the two cases are the reasons why (1) is true. In the world w' (1) is true because Dubya is the U.S. president, whereas in the world w'' (1) is true because Al Gore is the U.S. president. The reasons why (1) is true are different in the two worlds.

In the light of these considerations, a better candidate to perform the role of subject matter of a sentence S is the following: **why S is true**. To find out the subject matter of a sentence S , we must answer the correspondent question, i.e. *Why is S true?*

The definition of sentential subject matter provided by Yablo:

- (SS) The subject matter of S = the relation \mathbf{m} such that worlds are \mathbf{m} -dissimilar iff S is differently true in them.⁷²

The subject matter \mathbf{m} of a sentence S is the relation according to which different worlds are different with regard to \mathbf{m} if and only if the reasons for S to be true are different in the two worlds.

To represent the subject matter of a sentence S as **why S is true** that, in turn, is linked to the direct question *Why is S true?*, allows us to emphasise one more time that inclusion is not merely equivalent to implication. Let's consider again the example of conjunction and disjunction. To figure out whether a conjunct (p) is, in fact, part of the conjunction ($p \wedge q$), or not, the relation between them must be truth-preserving and aboutness-preserving. It is truth-preserving because conjunction implies each of its conjuncts. Is it also aboutness-preserving? We must consider their correspondent subject matters and we must check whether the subject matter of the conjunction includes that of the conjunct, or not. The subject matter of p is **why p is true** and the subject matter of $p \wedge q$ is **why $p \wedge q$ is true**. At this point we have to consider the correspondent direct question associated to the two subject matters and if the answer to the former is included in the answer to the latter, then we can conclude that the subject matter of the former is part of the subject matter of the latter and that conjunction is aboutness-preserving. In other words, we have to check whether the answer to the question *Why is p true?* is included in the answer to the question *Why is $p \wedge q$ true?*, or not. The result is that answering the latter means also answering the former: *Why is p true?* has as answer p , and *Why is $p \wedge q$ true?* has as answer p, q . Therefore, the latter

⁷² Ibid., p.41.

includes the former. We can conclude that in the case of conjunction the relation is aboutness-preserving and a conjunct is part of the conjunction.

What about disjunction? The proceeding is the same as before. We are wondering whether the disjunction ($p \vee q$) is part of the disjunct (p), or not. It is truth-preserving because the disjunct implies the disjunction, but is it also aboutness-preserving? To establish this we have to check whether the subject matter of $p \vee q$, i.e. **why $p \vee q$ is true**, is included in the subject matter of p , i.e. **why p is true**, or not. In the same way as before, we have to check whether the answer to the direct question associated to the former is included in the answer to the direct question associated to the latter, or not. In other words, we have to check whether the answer to *Why is $p \vee q$ true?* is included in the answer to the question *Why is p true?*, or not. Contrariwise to conjunction, in this case answering the latter question does not mean answering also the former: *Why is p true?* has as answer p , whereas *Why is $p \vee q$ true?* has two possible answers, either p , $\neg q$ or $\neg p$, q . Therefore the former does not include the latter. Disjunction, thus, is not aboutness-preserving because the disjunction is not part of its disjuncts.

If, on the one hand, the definition (SS) enables us to clarify the difference between implication and the relation of inclusion among hypotheses, on the other hand, it is not without its problems. (SS) seems not enough for an accurate definition of the subject matter of a sentence and it needs something more to be complete. If we limited the definition to what (SS) holds, three serious problems would arise.

The first problem is about the possibility for S to be false in a world w . According to (SS), if S is false in w , than things are in no way when S 's subject matter is concerned, because we have just defined the subject matter of S as **why S is true**, and the falsity of S is not taken into account. However, it is clear that S is false for a reason, even if this reason is not accounted by (SS).

The second problem with (SS) is that one should be able to understand what S is about while remaining ignorant of its truth-value. But this is not possible according to (SS) unless the subject matter of S is **why S is true**. In fact, if the subject matter of S is **why S is true**, then to know the subject matter of S , we must know first that S is true, but obviously it is not necessary to know the truth-value of a sentence in order to understand what it's about.

Lastly, the third problem is that the subject matter of a hypothesis and the subject matter of its negation must be the same, because the two sentences are clearly about the same things. However, (SS) doesn't allow for this.

All these problems seems to suggest that the issue actually amounts to just one: the subject matter of a sentence must tell not only the reasons why the sentence is true, but also the reasons why it is false. Hence, if s is the subject matter of the sentence S and it is defined as $s = \text{why } S \text{ is true}$, then \bar{s} is what Yablo calls the *subject anti-matter* of the sentence S and is defined as $\bar{s} = \text{why } S \text{ is false}$. It's now possible to define the *overall subject matter* \hat{s} of the sentence S as the set of S 's subject matter and S 's subject anti-matter, that is $\hat{s} = \{s, \bar{s}\}$.

At this point, it's easy to verify that the three problems are all solved thanks to the notion of overall subject matter of a sentence. In first place, the overall subject matter of S is defined wherever S has a truth-value and, for this reason, it is not necessary for S to be true in order to have a subject matter. In the second place, the overall subject matter of S is graspable even if we don't

know its truth-value. In fact, we don't need to know whether S is true or false in order to know its overall subject matter. Finally, as we wanted to be, the overall subject matter of S is identical to the overall subject matter of $\neg S$. This is very easily verifiable:

$$\hat{s} = \{s, \bar{s}\} = \{\neg\bar{s}, \neg s\} = \{\neg s, \neg\bar{s}\} = \neg\hat{s}.$$

Therefore, the concept of overall subject matter defined in this way seems to be exactly the notion of sentential subject matter we were looking for.

5.4.6 The Part of A about m

We have so far defined the notion of overall subject matter of a sentence. Remember, however, that Yablo's purpose is to say of certain false sentences that they are partially true. In order to achieve this goal we still didn't take into account an important step, namely we didn't find out yet a way to establish the part of a sentence A about a subject matter m .

Yablo makes a difference between two kinds of propositions, the thin and thick propositions. The *thin proposition* that A is A 's truth-conditional content, whereas the *thick* (or, *directed*) *proposition* that A is A 's truth-conditional content plus A 's subject matter. Therefore, the thick proposition that A is obtained by adding to the thin proposition that A , the subject matter of A . To use Yablo's symbolic notation we can define the thick proposition that A as follows:

- (DP) The directed proposition that A consists of:
1. $|A| = \mathbf{A}$ = A 's truth-conditional content;
 2. $\langle A \rangle = \mathbf{a}$ = A 's subject matter.

If the content of A is defined as a directed proposition, and if a directed proposition is defined as above, then it is clear that the part of A about m must be a directed proposition as well. To define the part of A about m (A_m), we must previously find out the directed proposition that A_m , that is, we must construct the following:

- (DP_m) The directed proposition that A_m consists of:
1. $|A_m| = A_m$'s truth-conditional content;
 2. $\langle A_m \rangle = A_m$'s subject matter.⁷³

Therefore, let's find out now the directed proposition that A_m . Let's consider the following sentence:

- (D) The number of atoms is constant over time.

With an utterance of D we don't want to convey anything about numbers, but rather we want to speak about atoms and concrete world in general. Hence, we must consider the part of D about the subject matter **concreta**. The thin proposition that D about **concreta** ($|D_{\text{concreta}}|$) is the set of worlds where D is

⁷³ Ibid., p.49.

true about the subject matter **concreta**. What should a world w be like, for D_{concreta} to be true in it? There are two possibilities for w : either w is a Platonist world in which D is true simpliciter, or w is a nominalist world in which D is false for reasons unrelated to the relevant subject matter, i.e. **concreta**.

But which is the part of D about **concreta**? According to Yablo's account, to be part of D , D_{concreta} needs a subject matter, and that subject matter has to be included in the subject matter of D . For what we previously said, this means that D_{concreta} needs ways of being true ($D_{\text{concreta}}\uparrow$) that are implied by D 's ways of being true ($D\uparrow$), and ways of being false ($D_{\text{concreta}}\downarrow$) that are implied by D 's ways of being false ($D\downarrow$).⁷⁴ Now, we have to answer the two corresponding direct questions. In the first place, we have to answer *How (or, why) is the part of D about **concreta** true?* If D is true in w because of $D\uparrow$, then the part of D about **concreta** is true in w because of the part of $D\uparrow$ being true about **concreta**. In the second place, we have to answer *How (or, why) is the part of D about **concreta** false?* Its falsmakers are going to be certain of D 's falsmaker. Which ones? Those which, in addition to not allowing D to be true, do not allow it to be true also about **concreta**.

In the light of these considerations, we can define the part of A about \mathbf{m} as follows:

- (P) The part of A about \mathbf{m} is the thick proposition that:
1. is true where A is true about \mathbf{m} ;
 2. has, for each $A\uparrow$, a truthmaker holding just where $A\uparrow$ is true about \mathbf{m} ;
 3. is false where A is false about \mathbf{m} .
 4. has, for each $A\downarrow$, $A\downarrow$ as a falsmaker just if $A\downarrow$ is not true about \mathbf{m} .⁷⁵

By applying this definition to the previous example we can conclude that the part of D about **concreta** is the proposition that is true in worlds with equally many atoms at all times, in virtue of facts like the following: there is always a single atom; there is always a pair of atoms; there are always three atoms, and so on. Similarly in order for the part of D about **concreta** to be false. It is the proposition that is false in worlds whose atoms become more or less plentiful, in virtue of facts like the following: there was an atom, and then there is a pair; there were no atoms, and then one appeared, and so on.⁷⁶

5.4.7 Yablo's Theory in a Nutshell

In this section we have shown the long path leading Yablo to the definition of the relation of aboutness. We have seen that the route followed by the philosopher is clear and careful enough not to skip even a single step. It can be summed up in the following schema that retraces the main steps of the argument:

⁷⁴ Of course, it's now clear that by "ways of being true" we mean truthmakers and by "ways of being false" we mean falsmakers.

⁷⁵ Ibid., p.54.

⁷⁶ Ibid., p.53.

- Semantic excuses;
- Partial truth;
- Parthood relation between hypotheses;
- Subject matter;
- Truth about a subject matter;
- Sentential subject matter;
- Aboutness;
- The part of A about **m**.

To briefly recap what we've learned so far, we have first brought some examples of semantic excuses, which are false sentences that can be used to express truths we cannot express otherwise. Then, we have defined this notion of partial truth by saying that a hypothesis is partially true if and only if it contains parts that are wholly true. At this point, we've moved to the definition of the relation of parthood between hypotheses and we have seen that it cannot be reduced to mere implication because it has an explanatory role that implication doesn't perform. On this line, the relation of parthood has been defined as the sum of implication and inclusion of subject matter, because it must be both truth-preserving and aboutness-preserving. More in details, B is part of A iff the inference from A to B is such that both A implies B and A's subject matter includes that of B. After having introduced the concept of subject matter we've defined it. The subject matter is a partition of logical space, or, in other words, a partition of propositions. In an intuitive way the subject matter can be meant as an indirect question. The subject matter **what you did last July** can be found out by answering the correspondent direct question: *What did you do last July?* We have, then, moved forward in the path and we have wondered what it means for something to be true about a subject matter. We've said that to say that S is true about **m** we need a world where **m** is kept unchanged and where S is true simpliciter. At the end, we have defined the notion of overall subject matter of a sentence S, namely, the sum of the subject matter and the subject anti-matter of S, where the subject matter of S is **why S is true** and the subject anti-matter of S is **why S is false**. Now, the relation of aboutness has been defined as the relation between a sentence and its subject matter. Lastly, Yablo has provided the definition of the part of a sentence A about the subject matter **m**, which involves the notions of truthmakers and falsemakers of a sentence.

What I'm going to show, now, is simply a rough sketch of how it is possible to take Yablo's account of aboutness one step forward in order to solve the problem of the liar paradox. Before that, I'm going to show that, despite Yablo has never explicitly defined his account of aboutness as a fictionalist theory, the two accounts can be translated into each other. In particular, it is possible to translate the key notions of aboutness' approach into fictionalist main concepts, and vice versa. Furthermore, I'll try to show that not only the aboutness' account has the great advantage of doing the same things fictionalism does, but also it does even more.

5.4.8 Tying Aboutness and Fictionalism

Besides technical details, the main point in Yablo's account is that it is possible to claim of certain sentences that they are false but partially true, that is, that they have a part that is wholly true. In particular, we saw that the wholly true part of a sentence is the one about a specific subject matter that, in fact, is the relevant subject matter for that sentence. If we make an appeal to Yablo's fictionalist vocabulary of his first works,⁷⁷ he will say that we must distinguish between two types of content of a sentence: the *semantic* (or *literal*) *content*, that is what makes certain sentences false; and the *real content*, that is what the speaker really wants to convey and what makes those sentences true.

One of the main problems in this distinction was raised by Amie Thomasson,⁷⁸ and can be summarized as follows: to make the distinction between literal and real content, the fictionalist must assign different truth-conditions to the two kinds of content; however, it's not clear in what the truth-conditions must differ and, in particular, it's not clear what the truth-conditions of the literal content add to those of the real content. As Thomasson puts it:

What more could it be supposed to take for the literal content to be literally true, than merely for the real content to be true? There might be some who are tempted to think that more *is* required for there to *really be* numbers, properties, or even marriages and corporations, than can follow via trivial transformations from the truth of the undisputed statement: namely, that there *really is some (new) object/individual* present. [Thomasson 2013, p.1039 (emphasis in original)]

According to Matteo Plebani,⁷⁹ it's possible to provide an answer to Thomasson's objection by using the concept of subject matter as defined by Yablo. As we saw, the content of a sentence is a directed proposition that is the sum of its truth-conditions and its subject matter. Since the subject matter is a partition of the set of worlds and since the truth-conditions of a sentence are defined as a set of worlds, and since a proposition is, by definition, a set of worlds, then it follows that the content of a sentence is a pair of sets of propositions. If the content of a sentence is defined not only on the basis of truth-conditions but also on the basis of its subject matter, we can explain the difference between real and literal content of a sentence without being forced to refer to truth-conditions and, so, by avoiding Thomasson's objection. In particular, as we already saw in the case of Hempel's raven paradox, a difference in the subject matter does not imply a difference in the truth-conditions of different sentences. The sentences "All ravens are black" and "All non-black things are non-ravens" have the same truth-conditions, because they are true under the same circumstances, but they have different subject matters: the former is about ravens and the latter is about the non-black things that are not ravens.

Let's try to clarify, now, Plebani's suggestion through a well-known example from applied mathematics:

⁷⁷ Yablo[2001 and 2005].

⁷⁸ Thomasson[2013].

⁷⁹ Plebani[2017].

(2) The Martian moons are two.

According to the fictionalist account of Yablo, (2) has two contents that must be kept separated – namely, the literal content implying an ontological commitment towards number two and numbers in general, and the real content that has nothing to say about numbers and it tells us how things are about Mars, its moons and the concrete world in general. Translated into the vocabulary of aboutness' account, this means that in the literal content of (2), besides its truth-conditions, there is a subject matter larger than the subject matter associated to its real content. The subject matter of the real content of (2) is **concreta**, because it is restricted to the concrete world, whereas the subject matter of the literal content of (2) includes also abstract entities and, hence, is larger than **concreta**.

Therefore, not only we can translate the fundamental concepts of Yablo's fictionalism into those of his own aboutness' approach, and vice versa, but it is also the case for the latter to be able to solve in a clear and elegant way the problems and objections raised against the former.

Furthermore, if we carefully think of Yablo's fictionalism and his aboutness' theory, we see that there is another connection between them. This connection is what enables us to translate what we can develop within one theory into the main concepts of the other, and vice versa. In other words, there's a direct connection between the relevant subject matter (aboutness' theory) and the relevant fiction (fictionalist account). In fact, it seems possible to directly move from the relevant subject matter of a sentence to the fiction of a fictionalist view, and vice versa. How are the two concepts tied? One turns out to be the complement of the other. Once we have the relevant subject matter of a sentence we can easily find the relevant fiction for that sentence, and vice versa.

I'm going now to elaborate this point through some examples from different regions of discourse. First, let's consider the sentence (2) from applied mathematics. We saw that, according to Yablo's aboutness theory, (2) is false but partially true, meaning that it is totally true about the relevant subject matter of the sentence, i.e. **concreta**. What about fictionalism? Fictionalists about mathematical discourse claim that (2) is literally false – that is, its literal content is false – but it's true if we interpret it in a fictionalist spirit – that is, if we consider its real content. The relevant fiction of (2) is the story of applied mathematics or, more in general, the story of abstract entities. In fact, according to the story of applied mathematics, (2) is true.

But the story of mathematics is nothing but the complement of the concrete world. In other words, if the subject matter and the fiction are added together we obtain the whole subject matter of the sentence and, in this specific case, if we add mathematics to the concrete world, we obtain the whole subject matter of (2). If we know, then, the relevant subject matter of a sentence, we can find out the correspondent fiction through a logical subtraction between the whole subject matter and the relevant subject matter of the sentence. In the light of these considerations, the equation is the following:

$$S = m + f$$

Where **S** is the whole subject matter of a sentence, **m** is its relevant subject matter and **f** is the correspondent fiction. To find out the relevant subject matter of a sentence when we know the correspondent fiction, we must simply subtract the fiction from the whole subject matter of the sentence ($\mathbf{m} = \mathbf{S} - \mathbf{f}$), and vice versa ($\mathbf{f} = \mathbf{S} - \mathbf{m}$).⁸⁰

Have a look now at another example from a different region of discourse:

- (3) Pegasus doesn't exist.

As we stressed in section 5.4.2, Yablo suggests to take the sentence (3) to be false but partially true, namely, true about the proper topic of the sentence. The relevant subject matter for the evaluation of (3) is **existent objects**: (3) is true if we consider how things are about existent things. In fact, it is true that Pegasus does not belong to the set of existent things, and this is exactly what an utterance of (3) tells us. On the other hand, the fictionalist analysis leads us to the conclusion that (3) is literally false but fictionalistically true. In other words, (3) is true according to the relevant fiction, i.e. the fiction of non-existent objects.

Let's now add the relevant subject matter of the sentence (3) to its relevant fiction. What do we obtain? If we add non-existent objects (fictionalism) to **existent objects** (aboutness) the result is the whole subject matter of (3). Also in this case, thus, the fiction stated by fictionalism can be interpreted as the complement of the relevant subject matter of the aboutness' account.

Finally, let's examine one last example we already mentioned, i.e. Sally's statement:

- (4) The tomato loses the property of being green and gains the property of being red.

As we said, if we are nominalists about properties, this sentence is false because there are no such things as properties. What does the analysis provided by the two accounts say? The theory of aboutness tells us that (4) is false but involves a part that is wholly true and that, by definition, makes the sentence partially true. The part wholly true is the one about the tomato, which doesn't say anything at all about properties. Hence, more in general, (4) is true about the relevant subject matter, i.e. **concreta**. The analysis provided by fictionalism, instead, is that (4) is in fact false, if taken at face value, but it is true according to the relevant fiction, i.e. the story of properties.

All these examples allow us to claim that the relevant subject matter of a sentence can be conceived as the complement of the relevant fiction for that sentence. Hence, the addition of these two components has as result the whole subject matter of the sentence without restriction of a sort.

⁸⁰ It's clear that **f** is a subject matter as well. In fact, what we obtain when we subtract a subject matter from another subject matter must be a subject matter as well. The choice not to use the bold font when we talk of fiction is to keep the notions belonging to aboutness' account (the subject matter) distinct from the notions belonging to fictionalism (the fiction).

For all these reasons, we can conclude that the theory of subject matter developed by Yablo can be seen as the translation in non-fictional terms of what fictionalism claimed.

However, there is a remark that is worth mentioning. As I already said at the end of the previous section, not only the aboutness' account has the great advantage of doing the same things fictionalism does, but also it does even more. It is true that the key notions of fictionalism can be translated into the key notions of aboutness' theory, but there are also certain decisive advantages in the latter over the former. First, aboutness' account is wider than fictionalism because it has a wider range of applications. Roughly speaking, fictionalism is a sectorial perspective, whereas aboutness' approach applies to every sentence in the language without restriction. In other words, in order to be fictionalists we must specify the region of discourse which we want to take fictionally, because, for instance, we can be fictionalists about mathematical discourse but not about possible worlds. Aboutness' account, instead, is wider and more generic because once you assume it, you can apply it indifferently to all sentences of all areas of discourse without restrictions. The notion of subject matter, hence, is a more generic notion than that of fiction. Secondly, contrary to fictionalism, aboutness' account transcends the issue of the right metaphysical theory to assume. In other words, as we have already pointed out, aboutness' theory is an agnostic view with regard to the metaphysical issues about the existence of the objects which the sentences we are taking into account refer to.⁸¹ In order to be fictionalist about mathematical discourse, one must first embrace a nominalistic view of mathematical objects. On the contrary, aboutness' theory makes a deal with both nominalists and Platonists in philosophy of mathematics and, more in general, with both realists and those assuming a thinner ontology. Whether you believe in the existence of properties or not does not matter to the evaluation, for instance, of sentence (4). Sentence (4) is true about **concreta** and about this subject matter the sentence is true whether you are nominalist or realist about properties. On the other hand, your ontological commitment does matter – and it matters a lot – in the case of fictionalism. In fact, if you believe in the existence of properties, then you don't need a fictionalist approach of properties-talk at all, and sentences involving properties-terms will be true *tout court*, i.e. literally true as well. The same applies, for instance, to the ontological issue in philosophy of mathematics as well. Aboutness theory is neutral with regard to the problem of whether we should assume a nominalist or a Platonist account of mathematical objects. According to both Platonists and nominalists, the sentences of applied mathematics are true when their relevant subject matter (i.e. **concreta**) is concerned. However, the same cannot be said about fictionalism. Whether you believe in the existence of mathematical entities or not is, indeed, a very important issue in the case of fictionalism. After all, you won't have the necessity of embracing a fictionalist view of mathematical discourse if you believe in the existence of mathematical objects. The same argument holds also for the other regions of discourse involving reference to problematic entities, such as propositions-talk, moral values-talk, and so on.

⁸¹ See footnote 58.

In conclusion, it's true that we can move from fictionalism to aboutness' theory at any time because each theory is translatable into the other, but it's also true that the latter has some great advantages over the former. If we use aboutness' theory, we can set the ontological issue aside but the same doesn't hold for fictionalism.⁸² This is why I found interesting this new account by Yablo and I wish to show that it also provides an adequate solution to the problem of the liar.

5.4.9 What *About* the Liar Paradox?

The aboutness' account developed by Yablo seems to be applicable to every region of discourse, meaning that there's nothing preventing us from interpreting every sentence in the way the philosopher described and this is, in fact, his main aim. What *about* semantic paradoxes? Is it possible to apply Yablo's account of subject matter also to the problem of semantic paradoxes and, in particular, of the liar?

It seems that we can and good evidence of this can be found in a forthcoming paper by Yablo himself, in which the philosopher proposes an application of his account to the problem of a different semantic pathology known as truth-teller.⁸³

The truth-teller is a particular sentence attributing truth to itself:

(K) (K) is true.

Take a look at its evaluation. If we suppose (K) to be true, then this is exactly what (K) says and, hence, it is true. If, instead, we assume that (K) is false, then we have to negate what (K) says and, for this reason, it turns out to be false. The particular construction of the sentence, involving self-reference, implies that any attempt of evaluating the sentence is grounded only on (K) itself. The truth-teller is not a paradoxical sentence but it's pretty clear that it shares with the liar a semantic pathology of some sort, meaning that also the truth-value of the truth-teller cannot be easily and definitely determined. The situation is what Kripke calls *ungroundness*, meaning that there's no sentence not-involving the truth predicate – and, thus, there is no fact in the world – on which to base our valuation of both the truth-teller and the liar.

Have a look now at the so-called truthfulness-teller, namely, a sentence claiming of itself that it always says the truth. It can be formalized as follows:

(H) Everything (H) says is true.

When we try to evaluate the sentence, we find out that it is exactly in the same situation as the truth-teller. If we suppose (H) to be true, then, for what it says,

⁸² However, this does not mean that Yablo changes his mind on the topic of fictionalism. On the contrary, the ontological neutrality of Yablo's aboutness theory is only a contingent feature. He is an aboutness theorist, but he also endorses a fictionalist perspective of mathematical discourse. Yablo is, indeed, neutral about the metaphysics of truthmakers and this is exactly why he is a fictionalist. In particular, fictionalism represents for Yablo a motivation for the development of his aboutness theory.

⁸³ Yablo[forthcoming].

it follows that everything (H) says is true and, hence, also this very thing is true, i.e. (H) itself. If we suppose (H) to be false, then (H) really is false because it describes itself as true. Hence, every evaluation we try to give for the sentence, it turns out to be self-supporting, as in the case of (K).

Both in the case of the truth-teller and in the case of the truthfulness-teller the evaluation of the sentence depends only on the sentence itself. The situation, in Kripkean terms, can be described as follows: (K) and (H) are *instable* sentences, meaning that they are true in certain fixed points and false in others. However, since there is no fact on which to base the evaluation, both (K) and (H) are not in the minimal fixed point. The only way to have them in another fixed point is by arbitrarily evaluating them – i.e. arbitrarily assigning them either to the extension or to the anti-extension of the truth predicate. If at the beginning of the construction we evaluate them as true – i.e. if we assign them to the extension of the truth predicate –, then we will obtain all the non-minimal fixed points in which (H) and (K) are true. If, on the contrary, at the beginning of the construction we evaluate them as false – i.e. if we assign them to the anti-extension of the truth predicate –, then we will obtain all the non-minimal fixed points in which (H) and (K) are false.⁸⁴

Despite their common behaviour, the truthfulness-teller seems to be truer than the truth-teller, or, in other words, it seems more grounded than the truth-teller. How can we explain this intuition? In the words of Yablo, the situation is the following:

K and \neg K are both *unconditionally possible* each holds in *some* fixed points above every factual ground. The Truthfulness-Teller is different in this respect. H can be true only in fixed points making Moore's other statements true.⁸⁵ H is only *conditionally* possible. The result \neg H of negating it is, however, unconditionally possible just like K; whatever the ground-level facts may be, we can consistently treat H as false by virtue of its own falsity. [Yablo forthcoming, p.12 of the manuscript (emphasis in original)]

The point, thus, is that in order to evaluate (K) we should look only at (K) itself, whereas in order for (H) to be true it's necessary that (H) itself is true but also everything said by (H) must be true as well. If (H) said that snow is white, then in order to evaluate (H) we must have a look also at the facts, and, in particular, we must look at how things are about snow. On the contrary, for (H) to be false it is enough that (H) itself is false, because for the conjunction to be false it is enough that one conjunct is false.

In Yablo's terminology, this means that despite (K) and (H) share several features that apparently make them closer, there actually is an element of their content that makes them different. This is why the truthfulness-teller looks more grounded than the truth-teller. This element of the content is the subject matter.

Let's consider, now, the subject matter **facts**. In the case of (K) a changing in facts does not change the truth-value of the sentence. In fact, even

⁸⁴ For a brief overview over Kripke's theory see section 4.3.2.

⁸⁵ Yablo assumes that Moore is the one saying "Everything I say is true". Then, Moore, here corresponds to the truthfulness-teller.

if the snow were green, the value of (K) would be the same, because its value depends only on (K) itself and on our arbitrary decision. In the case of (H) things are different. In fact, a changing in facts might change its truth-value because, for instance, if the snow were green, then (H) would be false. Therefore, the truth-teller and the truthfulness-teller are different because they have different subject matters: the subject matter of (H) is larger than the subject matter of (K), because it contains also the reference to certain facts.

Therefore, it was already in Yablo's mind the idea of applying its aboutness' account to sentences involving the truth predicate in order to solve certain philosophical puzzles due to particular semantic pathologies. Take a look, now, at the liar and at what we can say of the liar if we use Yablo's approach of subject matter. I wish to stress that this is intended to remain, here, only an outline for possible future developments.

The aim of Yablo in the development of his account of aboutness is, remember, to legitimate certain semantic excuses we make in order to say something false. In other words, sometimes the utterance of a falsity is excusable because this is the only way we have to convey a truth embedded in the falsehood. According to Yablo, certain truths cannot be expressed except as parts of a larger falsehood. If this is the only way we have to express a truth, then it seems to be a small price to pay for uttering a falsehood. I'm wondering here whether the liar can fall within this category of sentences, or not.

Let's consider the revenge liar:

(Q) (Q) is not true.

Assuming that I'm right when I say that the liar is an acceptable semantic excuse, what does this mean? If the liar is an acceptable semantic excuse, then the liar is false but partially true. By definition, to have a true part means for a sentence to have a part that is wholly true. The relation of parthood in Yablo's account is defined as truth-preserving and aboutness-preserving. This means that in order for a sentence to be included in another sentence, on the one hand, the former must be implied by the latter and, on the other hand, its subject matter must be included in the subject matter of the latter. In the specific case of (Q), there must be a subject matter included in the whole subject matter of (Q) that makes (Q) partially true. What's this subject matter?

Before we talk of the relevant subject matter for the liar, let's have a look at some examples from different regions of discourse. To start with, let's consider a sentence from mathematical-talk:

(5) The number of Dragons is 0.

According to Yablo's analysis, (5) is a false sentence because it refers to non-existent objects, namely, numbers. However, there's a sense according to which we wish to say that (5) is true. In fact, it seems that (5) includes a part that is totally true and that, for this reason, makes (5) partially true. What's the subject matter of the wholly true part of (5)? It's **concreta**. If we look at how things are when only concrete objects (and not mathematical objects) are involved, then we figure out that (5) is true because it is true that there are no dragons in the

concrete world. And, according to Yablo, this is really what we are talking about with an utterance of (5).

Consider now a sentence of a different kind:

(6) Pegasus is not in the room.

(6) is false because there is no object which the word “Pegasus” refers to. However, in virtue of our intuitions we wish to say of (6) that it is true, because it seems to us that, in fact, there is no such a thing as Pegasus in the room. Yablo says that the reason of this intuition is that the real topic of (6) is not Pegasus, but the room. In the light of this considerations, we can say that (6) includes a true part and that part is the one about the relevant subject matter we wish to convey with an utterance of (6), i.e. **this room**. If we look at how things are with regard to the room, then it is clear that (6) is true because it is true that in the room there is no such a thing as Pegasus.

Let’s now come back to talk of the liar. Things are not so different for it. What’s the relevant subject matter of a sentence such as the liar? In the previous section we saw that the whole subject matter of a sentence is obtained by the addition of the relevant subject matter and the correspondent fiction. Therefore, to find out the relevant subject matter of a sentence we must logically subtract the fiction from the whole subject matter.

The whole subject matters for the examples above can be schematized in the following way:

$$S_{(5)} = \begin{cases} m_{(5)} = \text{concreta} \\ f_{(5)} = \text{Platonic entities} \end{cases}$$

$$S_{(6)} = \begin{cases} m_{(6)} = \text{this room} \\ f_{(6)} = \text{Meinongian entities} \end{cases}$$

that, in the form of logical addition, can be stated also in the following way:

$$S_{(5)} = \text{concreta} + \text{Platonic entities}$$

$$S_{(6)} = \text{this room} + \text{Meinongian entities}$$

In the light of this, the whole subject matter of the sentence (5) turns out to be the set of relevant subject matter, i.e. **concreta**, and its relevant fiction, i.e. Platonic entities. In the case of (6) the whole subject matter is the set of relevant subject matter, i.e. **this room**, and its relevant fiction, i.e. Meinongian entities.

What about the revenge liar (Q)? We must, first, find out the relevant fiction associated to (Q) in the case of alethic fictionalism and, in particular in the case of fictionalism about semantic paradoxes. When we try to provide an evaluation of the liar and other paradoxical sentences, we are doing as we have the full T-schema without restrictions, that is, the instances of the T-schema for every sentence of the language, including defective sentences, like the liar. In other words, when we try to provide an evaluation of the liar, we are doing as we have the correspondent instance of the T-schema for the liar. Therefore, the

relevant fiction in the case of (Q) is the story of the unrestricted T-schema. Using the schemata we just introduced:

$$S_{(Q)} = \begin{cases} m_{(Q)} = ? \\ f_{(Q)} = \text{unrestricted T - schema} \end{cases}$$

or, in the form of logical addition:

$$S_{(Q)} = \text{unrestricted T-schema} + ?$$

In the light of the considerations so far made, to find the relevant subject matter of (Q) we need to find out something that added to unrestricted T-schema gives us the whole subject matter of the liar. That something is **grounded sentences**. The relevant subject matter for the liar is grounded sentences because the liar is true about that very subject matter, that is, the liar is true if we look at how things are with regard to grounded sentences, because it is true what the liar says about the set of grounded sentence, namely that is not in it.

A remark that is worth mentioning is that not only (Q) is true about **grounded sentences**, meaning that there's a world w' **grounded sentences**-equivalent to the actual in which (Q) is true simpliciter, but is also the case for (Q) to be false about the same subject matter, meaning that $(\neg Q)$ is true about **grounded sentences** as well. This means that there is also a world w'' **grounded sentences**-equivalent to the actual in which $(\neg Q)$ is true simpliciter. The liar, then, is a case of what Yablo calls *quasi-contradiction*, the phenomenon that both a hypothesis and its negation are true about the same subject matter. The phenomenon of quasi-contradiction can be formalized as follows:

$$A/m \ \& \ \neg A/m$$

where A/m is the part of A about the subject matter **m** and, similarly $\neg A/m$ is the part of $\neg A$ about the subject matter **m**. If we consider the part of A about **m** as the possibility that A, then we can translate in modal words the previous sentence as follows:

$$\diamond A \ \& \ \diamond \neg A$$

In other words, this means that there is a world w' accessible to w such that A is true in w' and there is a world w'' accessible to w such that $\neg A$ is true in w''.

If we apply what we just said with regard to the liar, what we obtain is that:

$$Q/\text{grounded sentences} \ \& \ \neg Q/\text{grounded sentences}$$

and, similarly, in modal terms:

$$\diamond Q \ \& \ \diamond \neg Q$$

In other words, starting from a world w , there's a world w' in which the subject matter **grounded sentences** is kept unchanged (i.e. w' is **grounded sentences**-equivalent to w) and in which the liar is true simpliciter, and there's a world w'' in which the subject matter **grounded sentences** is kept unchanged and the negation of the liar is true simpliciter. In modal words, this means that there's a world w' accessible to w in which the liar is true and there's another world w'' accessible to w in which the negation of the liar is true.

This doesn't amount to a commitment to the dialethic thesis that the liar is both true and false, or that both the liar and its negation are true. The dialethic thesis holds that both the liar and its negation are true in the actual world and this is completely different from Yablo's notion of quasi-contradiction, according to which there is a possible world where the liar is true and another possible world, different from the previous but in the same **grounded sentences**-cell, in which the negation of the liar is true. This is why Yablo calls this phenomenon quasi-contradiction.

As is well known, in modal logic it is not the case that the possibility of each conjunct amount to the possibility of the whole conjunction. In other words:

$$\neg(\diamond Q \ \& \ \diamond \neg Q) = \diamond(Q \ \& \ \neg Q)$$

It's possible for the apple to be on the table and it's possible for the same apple to be on the floor, but this does not amount to the possibility for the apple to be at the same time both on the table and on the floor. There's no possible world in which this is the case. In the vocabulary of subject matter's theory, everything we can say of the liar is what follows:

$$\exists w : w \in Q/\text{grounded sentences} \ \& \ w \in \neg Q/\text{grounded sentences}.$$

In other words, there is a world (the actual, indeed) such that the liar is true in this world about the subject matter **grounded sentences** and in that very world also the negation of the liar is true about the same subject matter **grounded sentences**. This means that partitioning the logical space, in our cell about **grounded sentences** (i.e. in the cell where the things about grounded sentences are in the same way as they are in the actual world) – i.e. in our **grounded sentences**-cell – both the liar and its negation are true. In other words, both the liar and its negation are in our **grounded sentences**-cell.

A final remark that is worth mentioning is about the relevant subject matter of the liar sentence. We said that the liar must be evaluated by keeping in mind that the real topic of the sentence is **grounded sentences**. Every attempt to put the liar inside the set of grounded sentences inexorably fails and pushes it on the outside. Another way in which we can understand the relevant subject matter of the liar is as **the minimal fixed point**. The grounded sentences are, by Kripke's definition, all sentences that are in the minimal fixed point, meaning that they are all and only the sentences receiving an evaluation in the minimal fixed point. For this reason, we can consider as relevant subject matter for the liar also **the minimal fixed point**, because it is equivalent to **grounded sentences**. In fact, the liar is false but partially true about **the**

minimal fixed point and the negation of the liar is false and partially true as well with regard to the same subject matter.⁸⁶

5.4.10 Conclusion

In this section I've tried to give an overview of a new approach that enriches traditional notions – i.e. partial truth, subject matter, aboutness, and so on – with a new and innovative meaning. After having shown the part of Yablo's theory that mostly interested my analysis, I've tried to do two different things: on the one hand, I've tried to emphasize why this account is an adequate choice and, in particular, I've tried to make clear how it's possible to translate the aboutness approach into fictionalism; on the other hand, I've tried to extend its applications in order to cover also the liar paradox. As I stressed again and again, the present work is intended to remain here only an outline for possible future developments.

Although it appears completely separated from Yablo's previous works, I've tried, first, to show how Yablo's aboutness theory is perfectly in accordance with them. In other words, I've tried to emphasise that the key notions of Yablo's aboutness account can be translated into the key notions of fictionalism, and vice versa. More in details, the analysis revealed that the relevant fiction for a fictionalist view can be viewed as the complement of the relevant subject matter of a sentence in the aboutness' theory. In this way, the whole subject matter of a sentence turns out to be the logical sum of the relevant subject matter and the correspondent fiction.

In the second place, I've tried to emphasise what the different areas of application of the theory are. Among those areas, I found interesting, in particular, the one about semantic paradoxes. If, in fact, the liar, in a fictionalist spirit, is true (and false) according to the fiction of the unrestricted T-schema, then it's easy to find out its relevant subject matter in order to apply Yablo's aboutness theory to it. The relevant subject matter for the liar turns out to be **grounded sentences** (or, alternatively **the minimal fixed point**), since the liar is true with regard to how things are about grounded sentences because what the liar says about the set of grounded sentence is true, namely that it does not belong to it.

But there's more to say on this point. Not only we can say of the liar that is true about **grounded sentences**, but we can also say of the negation of the liar that is true about the same subject matter. The reason is that there's a world w' **grounded sentences**-equivalent to the actual where the liar is true simpliciter and there's a world w'' **grounded sentences**-equivalent to the actual where the negation of the liar is true simpliciter. And this, by definition, means of the liar that it is partially true and that also its negation is partially true. We are dealing, therefore, with a paradigmatic example of the

⁸⁶ Actually, with regard to the liar we could take into account the more general subject matter **fixed point**, without any other specification on which fixed point, because the liar is in no subject matter. However, there are other non-paradoxical sentences that share with the liar the semantic pathology of ungroundness – such as, among those we considered, the truth-teller and the truthfulness-teller – and that are in certain fixed points different from the minimal. Therefore, to provide a uniform analysis of this sentences, it seems more appropriate to use as relevant subject matter **the minimal fixed point**.

phenomenon that Yablo calls quasi-contradiction and that, as we emphasised in several places, doesn't amount to a commitment to true contradictions, as the dialetheist conceived them.

I wish to stress one more time that what I've done in this section is not to be considered a complete work, but rather it's only a sketched attempt to provide an alternative solution to the liar and, at the same time, to emphasise the greatness of Yablo's account that can be applied to several, if not all, areas of discourse, included the one about semantic paradoxes. The attempt must be enriched in several ways, for instance, with the adding of more technical details, but I hope I've clarified the underling idea and the possible future developments that can be make.

Conclusion

In a paper published in 2004, J.C. Beall argues in favour of a perspective known as *constructive methodological deflationism*, which corresponds to the attempt to defend the idea that a deflationary and truth-theoretic fictionalist theory must adopt dialetheism in order to solve the problems due to the liar paradox.

Aim of this work was to figure out whether the union of these three theories really is a necessity, as Beall claimed, or it is possible to keep them apart without losing any benefit at all. In particular, the dissertation has developed through the analysis of the consequences of certain specific ways in which the three theories directly involved in Beall's approach can be combined. They correspond to three different areas about the truth predicate: dialetheism is taken to represent the logico-semantic area; deflationism corresponds to the so-called ontological area; and alethic fictionalism is representative of the area about the discourse. The possibility of union among the theories has never been questioned. The purpose of the work was to prove that the union of all the three theories and certain combinations among them is anything but necessary and it turned out to be merely possible.

The first part of the analysis has been focused on the fusion of the first two theories, namely, dialetheism and deflationism. The aim was to highlight that the two theories are far from being indispensable to each other. In the light of the works by Armour-Garb and Beall, the analysis evolved in the attempt to answer the following complementary questions: "should dialetheists be deflationists?" and "should deflationists really be dialetheists as argued by Armour-Garb and Beall?"

In order to provide an adequate answer to the first question a detailed comparison between the dialethic and non-deflationary theory of Priest and the dialethic and deflationary theory of Beall has been mandatory. Dialetheism is known for being a neutral perspective towards the issue of the nature of truth and the most important and known theories that address this matter. However, to not conflict with any theory of the nature of truth doesn't lead to the consequence that dialetheists cannot have any preference on the theory that is most appropriate to adopt. What I wanted to do is to show that actually there is no special feature in the theory of Priest that makes it better than the theory developed by Beall, and vice versa. More specifically, the differences between the two theories intersect two different areas of interest, namely, the metaphysical and logico-semantic areas. With regard to the metaphysical differences, the main point is about the different kind of true contradictions admitted by the accounts of the two philosophers. In the dialethic theory of Beall the dialetheias arise only at the logico-semantic level and not also in the true-free language, contrary to what is accounted by Priest's theory. Despite this metaphysical divergence, the most interesting differences for my analysis were the logico-semantic ones. The central point of disagreement between the two theories is the validity of the intersubstitutivity principle. We saw that the full transparency of the truth predicate – that is, the possibility of substituting A

Conclusion

with $T(A)$ and vice versa in every transparent context *salva veritate* – is a bounding requirement for every deflationary theory of truth. This characteristic of the truth predicate is indeed what allows the deflationist to take her metaphysical intuitions about the nature of truth into the logico-semantic structure of her account. As we stressed again and again, the only way the deflationary has to obtain the full transparency of the truth predicate is by means of the intersubstitutivity principle, which, we know, holds only in Beall's theory and not in Priest's one. The consequences of the validity of the intersubstitutivity principle for a theory of truth are twofold: on the one hand, all the gluts turns out to be also gaps; and, on the other hand, all internal contradictions give rise to an equal amount of correspondent external contradictions. However, both consequences were not decisive for the analysis. Provided that the intersubstitutivity principle has no awkward consequence for a dialethic theory, why does Priest prevent his theory from validating it? The reason is that, by using Priest's own words, "contradictions should not be multiplied beyond necessity"¹ and, according to Priest, the validity of the intersubstitutivity principle will lead his theory right to this unpleasant result. Nevertheless, that contradictions should not be multiplied beyond necessity depends from what one takes to be within the scope of the notion of necessity. We saw, indeed, that according to Beall the intersubstitutivity principle is a bounding requirement for his theory because its validity is the only way the philosopher has to obtain the full transparency of the truth predicate and the full transparency of the truth predicate, in turn, is a bounding requirement in order to move his deflationary intuitions about the nature of truth from the metaphysical to the logico-semantic level. Things are different according to Priest, who assumes a Dummettian perspective of the nature of truth that needs a thicker property of truth to be maintained. For this reason, the account developed by Priest doesn't need the intersubstitutivity principle in order to be coherent with his metaphysical ideas. Therefore, the following consequence is mandatory: on the one hand, from Priest's perspective the intersubstitutivity principle actually multiplies contradictions beyond necessity because it is not a *sine qua non* principle of his theory; on the other hand, Beall's metaphysical intuitions require the validity of the intersubstitutivity principle and, for this reason, the resulting proliferation of true contradictions is not beyond necessity. Therefore, if we look only at the logico-semantic structure of the two accounts, there seems to be no criterion helping us in the decision of which is the best theory between Priest's and Beall's one. The only way to make such a choice is by shifting from the logico-semantic to the metaphysical level. Since it's not possible to establish which is the best dialethic theory from a logico-semantic point of view, and since the only benchmark for the decision – namely, that contradictions should not be multiplied beyond necessity – strictly depends from what we take to be within the scope of necessity, and since what is necessary for a semantic theory depends from the theory about the nature of truth its advocates assume, then the metaphysical believes and the correspondent theories about the nature of truth will be the proper criterion at stake. Thus, wondering whether it is better the dialethic theory of Priest or the

¹ Priest[1987], p.71 and p.116.

Conclusion

one developed by Beall means nothing but wondering whether it is better the theory about the nature of truth subscribed by Priest or that embraced by Beall.

We saw that the question “should dialetheists be deflationists?” had a negative answer. What about the complementary question: “should deflationists really be dialetheists?” For this purpose, focus of the investigation became the comparison between the following two deflationary theories: Beall’s dialetheic theory and Field’s paracomplete one. The aim was to prove that if the two theories are good, then they are equally good and there’s no reason to prefer the dialetheic one over the other, contrary to what Armour-Garb and Beall maintained. In the first place, we saw what are the main points of similarity between the two theories. Since they are both deflationary theories, the intersubstitutivity principle holds in both Beall’s and Field’s account. Moreover, both theories are three-valued logics and the matrix of their truth tables is the same, the only difference is in the interpretation of the third value. The main divergence in the logic is about what classical rule they reject: the dialetheist rejects the rule of the *ex contradictione quodlibet*, whereas the paracompletist rejects the law of excluded middle. However, the consequences of this fundamental difference are of no help in establishing what is the best deflationary theory because they actually highlight the dual character of the theories. The main points of comparison between the two theories are essentially four: the solution to the curry’s paradox; the non-classical treatment of the notions of acceptance and rejection; the phenomenon of the revenge liar; and the problem of the expressive limitedness of the language. We saw that both theories face the problem of the curry’s paradox and the solutions provided by the two philosophers mirror each other: Beall rejects the rule of the elimination of the conditional, i.e. modus ponens; whereas Field rejects its counterpart, i.e. the rule of the introduction of the conditional. Similar remarks can be made if we look at pragmatics. Both Beall and Field reject the interdefinability of the notions of acceptance and rejection and the reasons they provide are different but dual. According to Beall it’s possible to accept both a sentence and its negation, whereas according to Field it’s possible to reject both of them. Moreover, both philosophers agree about the incompatibility of the two notions. The analysis also revealed that not only the two deflationary theories have specular consequences, but they also meet the same problems. On the one hand, both dialetheists and paracompletists can solve the problem of the revenge liar only on pain of the incompleteness of their languages. On the other hand, they can choose to keep their language expressively complete only on pain of inescapable contradictions.

In the light of these considerations, the conclusion is mandatory: there is no decisive advantage of a theory over the other. Therefore, to decide which one is the best between Beall’s and Field’s deflationary theory means making an arbitrary choice only on the basis of the following observations: on the one hand, we have a theory that keeps consistency as a bounding requirement but has the unpleasant result of a very complicated and entangled construction aimed towards an adequate conditional; on the other hand, we have a theory that has simplicity as a distinctive feature of its structure but, in contrast, is forced to admit inconsistency among its main theses. Therefore, the choice is between consistency – corresponding to Field’s paracomplete theory – and simplicity – corresponding to Beall’s dialetheism – and such a choice cannot be

Conclusion

based on nothing but an arbitrary preference on these aspects because, as we saw, from a logico-semantic perspective there is nothing allowing us to choose between the two theories.

In the second part of the dissertation the last area about the truth predicate, namely the area about the discourse, represented by truth-theoretic fictionalism, entered into the debate. The aim was to figure out whether the combination of all the three theories taken into account – namely, dialetheism, deflationism and fictionalism – is possible and perhaps desirable, and what are the difficulties we will meet if we choose to give rise to such a union.

The structure of the analysis of this second part perfectly mimics the structure of the first part. Even in this case, my purpose was to answer two complementary questions: “should dialetheists be fictionalists about truth-talk?” and “should alethic fictionalists be dialetheists?”

The attempt to answer the first question inevitably went through the comparison between the inconsistency theories of truth and dialetheism. My aim here was to highlight what are the main reasons which led some advocates of the inconsistency theories of truth to assume a fictionalist perspective of truth-talk and to argue that such reasons are not shared by dialetheists. In order to account for semantic paradoxes, the advocates of inconsistency theories of truth maintain the thesis that the language is intrinsically inconsistent. However, they don't aim to take this idea one step forward by admitting the existence of real true contradictions. Fictionalism comes in just at this level because it's what enables the advocates of inconsistency theories of truth to ontologically weaken the consequence that there are true contradictions. As a matter of fact, the theoreticians of inconsistency maintain that thanks to alethic fictionalism it is possible to claim that according to the fiction of truth there are true contradictions. This statement commits one only to the existence of dialetheias inside the fiction and not to the existence of true contradictions also outside it. If this is the main motivation for an advocate of inconsistency theories to assume alethic fictionalism, it is clear that dialetheist has no reason at all to embrace such a view of truth because dialetheism is, by definition, the theory that there are true contradictions. If fictionalism about a region of discourse has the virtue of avoiding an inconvenient ontological commitment towards the objects which the sentences of that region of discourse refer to, then fictionalism about truth-talk is of no help for a dialetheist, because dialetheists don't need to avoid the ontological commitment towards truth, even when it is the source of true contradictions. Therefore, the answer to the question “should dialetheists be fictionalists about truth-talk?” turned out negative. This doesn't mean that dialetheism is incompatible with truth-theoretic fictionalism, but that the union between the two theories is far from being mandatory for a dialetheist.

What was the answer to the second question about truth-theoretic fictionalism, i.e. “should alethic fictionalists be dialetheists?” According to Beall, fictionalists about truth-talk must necessary adopt a dialethic theory if they aim to solve the problem of the liar paradox. Nevertheless, the analysis showed that this is not the case. Indeed, the investigation proved that even if we admitted that dialetheism fits very well with the needs of a fictionalist perspective of truth-talk, it still wouldn't be the only theory to meet those fictionalist needs. More in details, the result of the analysis is that alethic

Conclusion

fictionalism is compatible with the solutions of the liar paradox provided by at least three other theories besides dialetheism. First, inconsistency theories of truth, because the fictionalist of truth-talk maintains that according to the fiction there are true contradictions but this does not lead automatically to the acceptance of true contradictions outside the fiction. Hence, we can keep the idea that the language is inconsistent without relying on dialetheism in virtue of inconsistency theories of truth. Secondly, the fictionalist automatically rejects the instance of the T-schema for the liar because it lacks the relevant prefix. Therefore, a perfectly classical solution to the liar is available to the alethic fictionalist. Lastly, the fictionalist about truth-talk can choose a meaningless strategy in order to solve the liar paradox. According to this latter strategy, the liar lacks of content. In particular, according to Armour-Garb and Woodbridge, the liar lacks of specifying the conditions necessary for it to have content and, hence, it is meaningless. In conclusion, dialetheism is far from being the only possibility for a fictionalist about truth-talk and we can claim that fictionalists about truth-talk should not be dialetheists. The reason is not the incompatibility between the two theories, but rather because there are other theories, beyond dialetheism, that are compatible with alethic fictionalism.

The result of the analysis allows me to conclude that, although there is never or almost never anything in the structure of dialetheism, deflationism and alethic fictionalism preventing their union, their combination is far from being necessary, contrary to what Beall claimed. It's always possible to bring the three theories together but there is no fundamental feature that forces this union. The choice of making such a union will be dictated by believes about semantics, metaphysics and the discourse about the truth predicate of everyone. If these believes fit well with what the three theories hold – in particular that there are true contradictions; that the truth predicate does not correspond to any substantive property of truth and is a mere device useful only for expressive purposes; and that the discourse about truth is a construction that works through elements of fiction –, then it will be possible to consider setting up an account that keep them together. If, on the contrary, this is not the case because, for instance, one shares the core ideas only of one of the theories and not the others, then she is not destined to suffer the failure of her own account, but rather still has certain appropriate alternatives which she can rely on and that turned out to be equally good for the purpose for which are called into question.

Finally, the dissertation ends with a new attempt of providing an alternative solution to the liar paradox based on the aboutness account developed by Yablo. The choice of the account is due on a twofold reason. On the one hand, the account developed by Yablo is perfectly in line with the discussion made so far because Yablo's fictionalist ideas in philosophy of mathematics are what move the philosopher in the development of such an aboutness approach. One of my aims, indeed, was to show that fictionalism and aboutness are strictly connected, meaning that it is always possible to translate the main concept of fictionalism into the key notions of aboutness, and vice versa. Therefore, it is possible to move in every moment from one theory to the other. On the other hand, aboutness account has the great advantage to do exactly what fictionalism does and even more. First, it is a wider theory than fictionalism because has a wider range of applications.

Conclusion

Roughly speaking, fictionalism is a sectorial perspective, whereas aboutness approach applies to every sentence in the language without restriction of a sort. In order to be fictionalists we must specify the region of discourse which we want to take fictionalistically because, for instance, we can be fictionalists about mathematical discourse but not about possible worlds. Aboutness account, instead, is wider and more generic because once you assume it, you can apply it indifferently to all sentences of all areas of discourse without restrictions. The notion of subject matter, hence, is a more generic notion than that of fiction. In the second place, contrary to fictionalism, aboutness account transcends the metaphysical issue, meaning that it is regardless to the theory of the nature of truth one aims to assume. In order to be fictionalist about mathematical-talk, one should embrace a nominalistic theory of mathematical objects. On the contrary, the aboutness approach is neutral with respect to the nominalists-Platonists dispute and in particular to every dispute that sees realists and those assuming a thinner ontology as rivals. In this regard, I showed that the sentence “the number of Martian moons is two” doesn’t create disagreement between Platonists and nominalists if we assume the aboutness account of Yablo, contrary to what historically happens. As a matter of fact, both nominalists and Platonists agree in saying that the sentence is true about its relevant subject matter, that is, **concrete world**. For all these reasons Yablo’s aboutness theory seems to be an adequate choice to be improved in order to account also for the liar paradox. In particular, aboutness account allows to say that both the liar and its negation are false but partially true sentences, namely, they have a true part that is that about the relevant subject matter for those sentences. The relevant subject matter for the liar and its negation is **grounded sentences** and, we saw, both sentences are true with regard to it because they both can be seen as saying of themselves that they don’t belong to the set of grounded sentences and this is plainly true. Yablo calls this phenomenon *quasi-contradiction* and it does not force us to admit the existence of true contradictions à la Priest. The reason is that in order to speak of true contradictions, the liar and its negation should be wholly true in the actual world, but in Yablo’s aboutness theory we commit ourselves only to their partial truth about the same subject matter in the actual world. In other words, both the liar and its negation are wholly true in two different possible worlds that are in the same **grounded sentences-cell** of the actual one. As I stressed again and again in the last chapter, this last part of the dissertation about the new approach to the liar paradox was destined to remain here only a sketched attempt, but I hope it resulted to be an interesting outline to be enriched in further investigations and future developments.

Bibliography

- Aristotle (1928). *Metaphysics*. Translated into English by W.D. Ross, in *The Works of Aristotle*, vol. VIII.
- Armour-Garb, B. (2001). Deflationism and the Meaningless Strategy. *Analysis*, **61** (4), pp.280-289.
- Armour-Garb, B. (2004). Minimalism, the Generalization Problem and the Liar. *Synthese*, **139** (3), pp.491-512.
- Armour-Garb, B. (2012). Deflationism (about Theories of Truth). *Philosophy Compass*, **7** (4), pp.267-277.
- Armour-Garb, B. (2013). A Minimalist Theory of Truth. *Metaphilosophy*, **44** (1-2), pp.53-57.
- Armour-Garb, B. and Beall, J.C. (2001). Can Deflationists Be Dialetheists? *Journal of Philosophical Logic*, **30** (6), pp.593-608.
- Armour-Garb, B. and Beall, J.C. (2002). Further Remarks on Truth and Contradiction. *The Philosophical Quarterly*, **52** (207), pp.217-225.
- Armour-Garb, B. and Beall, J.C. (2003a). Minimalism and the Dialetheic Challenge. *Australasian Journal of Philosophy*, **81** (3), pp.383-401.
- Armour-Garb, B. and Beall, J.C. (2003b). Should Deflationists be Dialetheists? *Noûs*, **37** (2), pp.303-324.
- Armour-Garb, B. and Beall, J.C. (eds.) (2005). *Deflationary Truth*. Chicago and La Salle: Open Court Press.
- Armour-Garb, B. and Woodbridge, J.A. (2014). From Mathematical Fictionalism to Truth-Theoretic Fictionalism. *Philosophy and Phenomenological Research*, **88** (1), pp.93-118.
- Armour-Garb, B. and Woodbridge, J.A. (2015a). *Pretence and Pathology: Philosophical Fictionalism and its Applications*. Cambridge: Cambridge University Press.
- Armour-Garb, B. and Woodbridge, J.A. (2015b). Truth, Pretence and the Liar Paradox. In T. Achourioti, H. Galinon, J. Martínez Fernández and K. Fujimoto (eds.), *Unifying the Philosophy of Truth*, volume 36 of the series Logic, Epistemology, and the Unity of Science, Springer, pp.339-354.
- Azzouni, J. (2003). The Strengthened Liar, the Expressive Strength of Natural Languages, and Regimentation. *The Philosophical Forum*, **34**, pp.329-350.

Bibliography

- Balaguer, M. (2013). Fictionalism in the Philosophy of Mathematics. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Fall 2013 Edition. <http://plato.stanford.edu/archives/fall2013/entries/fictionalism-mathematics/>.
- Beall, J.C. (2001). A Neglected Deflationist Approach to the Liar. *Analysis*, **61** (270), pp.126-129.
- Beall, J.C. (2002). Deflationism and Gaps: Untying “Not”s in the Debate. *Analysis*, **62** (276), pp.299-305.
- Beall, J.C. (2004). True and False – As If. In G. Priest, J.C. Beall and B. Armour-Garb (eds.), *The Law of Non-Contradiction: New Philosophical Essays*, New York: Oxford University Press, pp.197-215.
- Beall, J.C. (2005). Transparent Disquotationalism. In J.C. Beall and B. Armour-Garb (eds.), *Deflationism and Paradox*, Oxford: Clarendon Press, pp.7-22.
- Beall, J.C. (ed.) (2007). *Revenge of the Liar: New Essays on the Paradox*. New York: Oxford University Press.
- Beall, J.C. (2009). *Spandrels of Truth*. Oxford: Oxford University Press.
- Beall, J.C. (2013). Curry’s Paradox. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2013 Edition. <http://plato.stanford.edu/archives/spr2013/entries/curry-paradox/>.
- Beall, J.C. and Armour-Garb, B. (eds.) (2005). *Deflationism and Paradox*, Oxford: Clarendon Press.
- Beall, J.C. and Glanzberg, M. (2011). Liar Paradox. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Fall 2014 Edition. <http://plato.stanford.edu/archives/fall2014/entries/liar-paradox/>.
- Beall, J.C. and Priest, G. (2007). Not So Deep Inconsistency: a Reply to Eklund. *Australasian Journal of Logic*, **5**, pp.74-84.
- Berti, E. (2014). Objections to Aristotle’s Defence of the Principle of Non-Contradiction. In E. Ficara (ed.), *Contradictions: Logic, History, Actuality*, Berlin: De Gruyter, pp.97-108.
- Berto, F. (2006). *Teorie dell’assurdo: i rivali del principio di non-contraddizione*. Roma: Carocci Editore.
- Berto, F. (2007a). Non dire non! (Una proposta che Priest non potrà rifiutare). In F. Berto, *Scenari dell’impossibile*, Padova: Il Poligrafo, pp.25-61.
- Berto, F. (2007). *How to Sell a Contradiction: the Logic and Metaphysics of Inconsistency*. London: College Publications.

Bibliography

- Berto, F. (2008). Adynaton and Material Exclusion. *Australasian Journal of Philosophy*, **86**, pp.165-190.
- Berto, F. (2014). Absolute Contradiction, Dialetheism and Revenge. *The Review of Symbolic Logic*, **7** (2), pp.193-207.
- Berto, F. (2015). A Modality Called “Negation”. *Mind*, **124** (495), pp.761-793.
- Berto, F. and Bottai, L. (2015). *Che cos'è una contraddizione*. Roma: Carocci editore.
- Berto, F. and Plebani M. (2015). *Ontology and Metaontology. A Contemporary Guide*. London: Bloomsbury.
- Burge, T. (1979). Semantical Paradox. *Journal of Philosophy*, **76**, pp.169-198.
- Burgess, A. (2007). *Identifying Fact and Fiction*. Doctoral dissertation. Princeton University.
- Burgess, A. (forthcoming). *Truth in Fictionalism*. In M. Glanzberg (ed.), *Oxford Handbook of Truth*, Oxford: Oxford University Press.
- Burgess, A. and Burgess, J. (2011). *Truth*. Princeton: Princeton University Press.
- Burgess, J. (1983). Why I am Not a Nominalist. *Notre Dame Journal of Formal Logic*, **24**, pp.93-105.
- Burgess, J. (2004). Mathematics and Bleak House. *Philosophia Mathematica*, **12** (1), pp.18-36.
- Burgess, J. and Rosen, G. (1997). *A Subject with no Object*. Oxford: Oxford University Press.
- Chihara, C. (1979). The Semantic Paradoxes: A Diagnostic Investigation. *Philosophical Review*, **88**, pp.590-618.
- Colyvan, M. (2011). Fictionalism in the Philosophy of Mathematics. In E.J. Craig (ed.), *Routledge Encyclopedia of Philosophy*, Online edition. <https://www.rep.routledge.com/articles/fictionalism-in-the-philosophy-of-mathematics>.
- Dummett, M. (1959). Truth. *Proceedings of the Aristotelian Society*, New Series, **59**, pp.141-162.
- Eklund, M. (2002a). Deep Inconsistency. *Australasian Journal of Philosophy*, **80**, pp.321-331.
- Eklund, M. (2002b). Inconsistent Languages. *Philosophy and Phenomenological Research*, **64**, pp.251-275.
- Eklund, M. (2005). Fiction, Indifference and Ontology. *Philosophy and Phenomenological Research*, **71**, pp.557-579.

Bibliography

- Eklund, M. (2008). Reply to Beall and Priest. *Australasian Journal of Philosophy*, 6, pp.94-106.
- Eklund, M. (2015). Fictionalism. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Winter 2015 Edition. <http://plato.stanford.edu/archives/win2015/entries/fictionalism/>.
- Field, H. (1980). *Science Without Numbers: A Defense of Nominalism*. Princeton: Princeton University Press.
- Field, H. (1986). The Deflationary Conception of Truth. In G. MacDonald and C. Wright (eds.), *Fact, Science and Morality*, Oxford: Blackwell.
- Field, H. (1989). *Realism, Mathematics and Modality*. Oxford: Basil Blackwell.
- Field, H. (1992). Critical Notice: Paul Horwich's *Truth*. *Philosophy of Science*, 59 (2), pp.321-330.
- Field, H. (1994a). Deflationist Views of Meaning and Content. *Mind*, **103** (411), pp.249-184.
- Field, H. (1994b). Disquotational Truth and Facually Defective Discourse. *Philosophical Review*, **103** (3), pp.405-452
- Field, H. (1999). Deflating the Conservativeness Argument. *Journal of Philosophy*, **96**, pp.533-540.
- Field, H. (2002). Saving the Truth Schema from Paradox. *Journal of Philosophical Logic*, **31**(1), pp.1-27.
- Field, H. (2003a). A Revenge-Immune Solution to the Semantic Paradoxes. *Journal of Philosophical Logic*, **32**, pp.139-177.
- Field, H. (2003b). The Semantic Paradoxes and the Paradoxes of Vagueness. In J.C. Beall (ed.), *Liars and Heaps: New Essay on Paradox*, New York: Oxford University Press, pp.262-311.
- Field, H. (2005). Is the Liar Sentence Both True and False? In J.C. Beall and B. Armour-Garb (eds.), *Deflationism and Paradox*, Oxford: Clarendon Press, pp.23-40.
- Field, H. (2006). Review of Graham Priest, *Doubt Truth to Be a Liar*. *Notre Dame Philosophical Reviews*, **3**.
- Field, H. (2007). Solving the Paradox, Escaping Revenge. In J.C. Beall (ed.), *Revenge of the Liar: New Essays on the Paradox*, New York: Oxford University Press, pp.78-144.
- Field, H. (2008). *Saving Truth from Paradox*. Oxford: Oxford University Press.

Bibliography

- Field, H. (2010). *Precis of Saving Truth from Paradox*. *Philosophical Studies*, **147** (3), pp.415-420.
- Field, H. (2010). Replies to Commentators on *Saving Truth from Paradox*. *Philosophical Studies*, **147** (3), pp.457-470.
- Frege, G. (1948). Sense and Reference. *The Philosophical Review*, **57** (3), pp.209-230.
- Frege, G. (1958). The Thought: A Logical Inquiry. *Mind*, **65** (259), pp.289-311.
- Glanzberg, M. (2004). A Contextual-Hierarchical Approach to Truth and the Liar Paradox. *Journal of Philosophical Logic*, **33**, pp.27-88.
- Grover, D. (1992). *A Prosentential Theory of Truth*. Princeton: Princeton University Press.
- Gupta, A. (1993). A Critique of Deflationism. *Philosophical Topics*, **21** (1), pp. 57-81.
- Gupta, A. (2005). Do the Paradoxes Pose a Special Problem for Deflationism? In J.C. Beall and B. Armour-Garb (eds.), *Deflationism and Paradox*, Oxford: Clarendon Press, pp.133-147.
- Hodges, W. (2014). Tarski's Truth Definitions. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Fall 2014 Edition. <http://plato.stanford.edu/archives/fall2014/entries/tarski-truth/>.
- Horsten, L. (1995). The Semantical Paradoxes, the Neutrality of Truth and the Neutrality of the Minimalist Theory of Truth. In P. Cortois (), *The Many Problems of Realism*, Vol. 3 of *Studies in the General Philosophy of Science*, Tilburg: Tilburg University Press, pp.173-187.
- Horsten, L. (2009). Levity. *Mind*, 118, pp.555-581.
- Horwich, P. (1998). *Truth*. Oxford: Blackwell.
- Horwich, P. (2008). *Varieties of Deflationism*. *Philosophical Topics*, **36** (2), pp.29-43.
- Kalderon, M. (eds.) (2005). *Fictionalism in Metaphysics*. Oxford: Clarendon Press.
- Kleene, S.C. (1952). *Introduction to Metamathematics*
- Kripke, S. (1975). Outline of a Theory of Truth. *Journal of Philosophy*, **72**, pp.690-716.
- Lewis, D. (1978). Truth in Fiction. *American Philosophical Quarterly*, **15**, pp.37-46.
- Lewis, D. (1988). Statements Partly about Observation. *Philosophical Papers*, **17** (1), pp.1-31.

Bibliography

- Lewis, D. (1993). Mathematics in Megethology. *Philosophia Mathematica*, **1** (1), pp.3-23.
- Liggins, D. (2008). Nihilism Without Self-Contradiction. In R. Le Poidevin (ed.), *Metaphysical Grounding: Understanding the Structure of Reality*, Cambridge: Cambridge Univeristy Press, pp.177-196.
- Liggins, D. (2014). Constructive Methodological Deflationism, Dialetheism and The Liar. *Analysis*, **74** (4), pp.566-574.
- Lindström, S. (2001). Horwich's Minimalist Conception of Truth: Some Logical Difficulties. *Logic and Logical Philosophy*, **9**, pp.161-181.
- Littmann, G. and Simmons, K. (2004). A Critique of Dialetheism. In G. Priest, J.C. Beall and B. Armour-Garb (eds.), *The Law of Non-Contradiction: New Philosophical Essays*, New York: Oxford University Press, pp.314-335.
- Ludwig, K. (2002). What is the Role of a Truth Theory in a Meaning Theory? In J.K. Campbell, M. O'Rourke and D. Shier (eds.), *Meaning and Truth: Investigations in Philosophical Semantics*, New York: Seven Bridges Press, pp.142-163.
- Lukasiewicz, J. (1951). *Aristotle's Syllogistic: From the Standpoint of Modern Formal Logic*. Oxford: Oxford University Press.
- McGrath, M. (2000). *Between Deflationism & Correspondence Theory*. New York: Garland Publishing, Inc.
- Meinong, A. (1960). On The Theory of Objects. In R. Chisholm (ed.), *Realism and the Background of Phenomenology*, Glencoe, Ill.: Free Press, pp.76-117.
- Murzi, J. and Carrara, M. (2014). Denial and Disagreement. *Topoi*, **1**, pp.1-11.
- Parsons, C. (1974). The Liar Paradox. *Journal of Philosophical Logic*, **3**, pp.381-412.
- Plebani, M. (2017). Fictionalism vs Deflationism. A New Look. *Philosophical studies*. DOI 10.1007/s11098-017-0868-7
- Priest, G. (1979). The Logic of Paradox. *Journal of Philosophical Logic*, **8** (1), pp.219-241.
- Priest, G. (1987). *In Contradiction: A Study of the Transconsistent*. Dordrecht: Martinus Nijhoff. 2nd expanded edition, Oxford: Oxford University Press, 2006.
- Priest, G. (1989). Reductio ad Absurdum et Modus Tollendo Ponens. In G. Priest, R. Routley and J. Norman (eds.), *Paraconsistent Logic: Essay on The Inconsistent*, Wien: Philosophia Verlag, pp.613-626.
- Priest, G. (1992). What is a Non-Normal World? *Logique & Analyse*, **139-140**, pp.291-302.

Bibliography

- Priest, G. (1998a). To Be and Not To Be – That is the Answer. On Aristotle on the Law of Non-Contradiction. *Philosophiegeschichte und Logische Analyse*, **1998** (1), pp.91-130.
- Priest, G. (1998b). What's So Bad About Contradictions? *Journal of Philosophy*, **95**, pp.410-426.
- Priest, G. (2000). Truth and Contradiction. *The Philosophical Quarterly*, **50** (200), pp.305-319.
- Priest, G. (2001). Why It's Irrational to Believe in Consistency. In B. Brogaard and B. Smith (eds.), *Rationality and Irrationality; Proc. 23rd International Wittgenstein Symposium*, pp.284-293.
- Priest, G. (1995). *Beyond the Limits of Thought*. Cambridge: Cambridge University Press, 2nd expanded edition, Oxford: Oxford University Press, 2002.
- Priest, G. (2005). Spiking the Field-Artillery. In J.C. Beall and B. Armour-Garb (eds.), *Deflationism and Paradox*, Oxford: Clarendon Press, pp.41-52.
- Priest, G. (2006). *Doubt Truth to Be a Liar*. New York: Oxford University Press.
- Priest, G. (2007). Revenge, Field and ZF. In J.C. Beall (ed.), *Revenge of the Liar: New Essays on the Paradox*, New York: Oxford University Press, pp.225-233.
- Priest, G. (2010). Hopes Fade for Saving Truth. *Philosophy*, **85** (1), pp.109-140.
- Priest, G., Beall, J.C. and Armour-Garb, B. (eds.) (2004). *The Law of Non-Contradiction: New Philosophical Essays*. New York: Oxford University Press.
- Priest, G. and Berto F. (2013). Dialetheism. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Summer 2013 Edition. <http://plato.stanford.edu/archives/sum2013/entries/dialetheism/>.
- Priest, G., Routley R. and Norman J. (eds.) (1989). *Paraconsistent Logic: Essays on the Inconsistent*. München: Philosophia Verlag.
- Priest, G., Tanaka, K. And Weber, Z. (2015). Paraconsistent Logic. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Spring 2015 Edition. <http://plato.stanford.edu/archives/spr2015/entries/logic-paraconsistent/>.
- Quesada, M. (1976). Heterodox Logic and the Problem of the Unity of Logic (abstract). *Journal of Symbolic Logic*, **43**, p.354.
- Quine, W.V. (1956). Quantifiers and Propositional Attitudes. *The Journal of Philosophy*, **53**, pp.177-187.
- Quine, W.V. (1970). *Philosophy of Logic*. Englewood Cliffs, New Jersey: Prentice Hall.
- Quine, W.V. (1987). *Quiddities: An Intermittently Philosophical Dictionary*. Cambridge, MA: The Belknap Press of Harvard University Press.

Bibliography

- Quine, W.V. (1992), *Pursuit of Truth*. Harvard University Press.
- Ramsey, F.P. (1927). Facts and Propositions. *Proceedings of the Aristotelian Society*, 7 (1), pp. 153-170.
- Ramsey, F.P. (1931). *The Foundation of Mathematics and Others Logical Essays*.
- Rayo, A. and Welch P.D. (2007). Field on Revenge. In J.C. Beall (ed.), *Revenge of the Liar: New Essays on the Paradox*, New York: Oxford University Press, pp.234-249.
- Ripley, D. (2014). Embedding Denial. In C. Caret and O. Hjortland (eds.), *Foundations of Logical Consequences*, New York: Oxford University Press, pp.289-309.
- Sainsbury, R.M. (2009). *Fiction and Fictionalism*. London and New York: Routledge.
- Scharp, K. (2013). *Replacing Truth*. Oxford: Oxford University Press.
- Schwartz, J.F. (2014). *Quine, Disquotatation and Truth*. Doctoral dissertation, The University of Chicago.
- Shapiro, S. (2004). Simple Truth, Contradiction, and Consistency. In G. Priest, J.C. Beall and B. Armour-Garb (eds.), *The Law of Non-Contradiction: New Philosophical Essays*, New York: Oxford University Press, pp.336-354.
- Simmons, K. (1993). *Universality and the Liar*. Cambridge: Cambridge University Press.
- Simmons, K. (1999). Deflationary Truth and the Liar. *Journal of Philosophical Logic*, 28 (5), pp.455-488.
- Slater, B.H. (1995). Paraconsistent Logics? *Journal of Philosophical Logic*, 24 (4), pp.451-454.
- Soames, S. (1997). The Truth about Deflationism. *Philosophical Issues*, 8, pp.1-44.
- Stanley, J. (2001). Hermeneutic Fictionalism. *Midwest Studies in Philosophy*, 25 (1), pp.36-71.
- Stoljar, D. and Damnjanovic, N. (2014). The Deflationary Theory of Truth. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Fall 2014 Edition. <http://plato.stanford.edu/archives/fall2014/entries/truth-deflationary/>.
- Strawson, P.F. (1949). Truth. *Analysis*, 9, pp.83-97.
- Tarski, A. (1944). The Semantic Conception of Truth and the Foundations of Semantics. *Philosophy and Phenomenological Research*, 4, pp.341–376.

Bibliography

- Tarski, A. (1983). The Concept of Truth in Formalized Languages. In J.H. Corcoran (ed.), *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, 2nd edition, Indianapolis: Hackett Publishing Company. English translation by J.H. Woodger of “Der Wahrheitsbegriff in den formalisierten Sprachen”, *Studia Philosophica*, 1 (1935).
- Thomasson, A. (2013). Fictionalism versus Deflationism. *Mind*, **122** (488), pp.1023-1051.
- Van Fraassen, B. (1980). *The Scientific Image*. Oxford: Oxford University Press.
- Varzi, A. (2013). Fictionalism in Ontology. In C. Barbero, M. Ferraris and A. Voltolini (eds.), *From Fictionalism to Realism*, Newcastle: Cambridge Scholars Publishing, pp.133-151.
- Walton, K. (1990). *Mimesis as Make-Believe*. Cambridge: Harvard University Press.
- Walton, K. (1993). Metaphor and Prop Oriented Make-Believe. *European Journal of Philosophy*, **1**, pp.39-57.
- Woodbridge, J.A. (2003). Deflationism and the Generalization Problem. *Logica Yearbook 2003*, pp.285-297.
- Woodbridge, J.A (2005). Truth as a Pretence. In M. Kalderon (ed.), *Fictionalism in Metaphysics*, Oxford: Oxford University Press, pp.134-177.
- Yablo, S. (2000). A Paradox of Existence. In A. Everett and T. Hofweber (eds.), *Empty Names, Fiction and The Puzzle of Non-Existence*, Stanford: CSLI Publications, pp.275-312.
- Yablo, S. (2001). Go Figure. A Path Through Fictionalism. In P. French and H. Wettstein (eds.), *Midwest Studies in Philosophy Volume XXV: Figurative Language*, Oxford: Blackwell, pp.72-102.
- Yablo, S. (2002). Abstract Objects: A Case Study. *Philosophical Issues*, **12**, pp.220-240.
- Yablo, S. (2003). New Grounds for Naive Truth Theory. In J.C. Beall (ed.), *Liars and Heaps: New Essay on Paradox*, New York: Oxford University Press, pp.312-330.
- Yablo, S. (2005). The Myth of The Seven. In M. Kalderon (ed.), *Fictionalism in Metaphysics*, Oxford: Oxford University Press, pp.88-115.
- Yablo, S. (2014). *Aboutness*. Princeton: Princeton University Press.
- Yablo, S. (forthcoming). Knights, Knaves, Truth, Truthfulness, Grounding, Tethering, Aboutness and Paradox. In M. Fitting (ed.), *Essay for Raymond Smullyan*.