Philosophical Semantics

Philosophical Semantics:

Reintegrating Theoretical Philosophy

By

Claudio Costa

Cambridge
Scholars
Publishing

Philosophical Semantics: Reintegrating Theoretical Philosophy

By Claudio Costa

This book first published 2018

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Copyright © 2018 by Claudio Costa

Cover: Edward Robert Hughes: *The Valkyrie's Vigil*

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-5275-0865-X
ISBN (13): 978-1-5275-0865-1

TABLE OF CONTENTS

Preface xi

Acknowledgments xviii

Chapter I 1
Introduction
1. Ernst Tugendhat's analysis of singular predicative statements
2. The virtue of comprehensiveness

Appendix to Chapter I 8
How Do Proper Names Really Work? (Cutting the Gordian Knot)
1. A meta-descriptivist rule for proper names
2. Identification rules at work
3. Objection of vagueness
4. Signification
5. Ignorance and error
6. Rigidity
7. Rules changeability
8. Names versus descriptions
9. Autonomous definite descriptions
10. Kripke's counterexamples
11. Donnellan's counterexamples
12. Explanatory failure of the causal-historical view

Chapter II 42
The Most Suitable Methodology for Conceptual Analysis
1. Common sense and meaning
2. Ambitious versus modest common sense
3. Resisting changes in worldviews
4. Primacy of established knowledge
5. Philosophizing by examples
6. Tacit knowledge of meaning: traditional explanation
7. A very simple model of semantic-cognitive rule
8. Criteria versus symptoms
9. Challenges to the traditional explanation (i): John McDowell

10. Challenges to the traditional explanation (ii): Gareth Evans
11. Unreflected semantic cognitions
12. Conclusion

Appendix to Chapter II 79
Modal Illusions: Against Supra-Epistemic Metaphysical Identities
Addendum: disposing of externalism

Chapter III 115
Wittgensteinian Semantics
1. Semantic-cognitive link
2. Why cannot reference be meaning?
3. Failure of Russell's atomistic referentialism
4. Meaning as a function of use
5. Meaning as a kind of rule
6. Meaning as associations of rules
7. Meaning and language-games
8. Meaning and form of life
9. Tying the threads together
10. Criteria and symptoms revisited
11. Transgressions of the internal limits of language
12. The form of semantic-cognitive rules
13. What is wrong with the private language argument?
14. Concluding remarks

Appendix to Chapter III 154
Trope Theory and the Unsustainable Lightness of Being
1. Introducing tropes
2. Tropes and universals
3. Tropes and concrete particulars
4. Formal tropes
5. Conclusion

Chapter IV 178
An Extravagant Reading of Fregean Semantics
1. Reference of a singular term
2. Sense of a singular term
3. Reference of a predicative expression
4. Ontological level
5. Referring to particularized properties (tropes)
6. Difficulties with the concept of unsaturation

7. Unsaturation as ontological dependence
8. Sense of a predicative term
9. The dependence of the predicative sense
10. The concept of horse paradox
11. Existence as a property of concepts
12. Existence as a property of conceptual rules
13. Two naïve objections
14. Attributing existence to objects
15. The existence of objects and its identification rules
16. Existence of spatiotemporal locations: indexicals
17. Advantages of the higher-order view of existence
18. Ubiquity of existence
19. Answering some final objections
20. Reference again: a metaphysical excurse (Mill)
21. The reference of a sentence as its truth-value
22. Logical structure of facts
23. Ontological nature of facts
24. Church's slingshot argument
25. Sub-facts and grounding facts
26. Taking seriously the sentence's reference as a fact
27. The riddle of identity in difference
28. Contexts of interest: no need for a necessary a posteriori
29. Sense of sentences: the thought
30. The thought as the truth-bearer
31. Facts as true thoughts?
32. The thought as a verifiability rule
33. Frege's Platonism
34. Avoiding Frege's Platonism
35. Further ontological consequences
36. A short digression on contingent futures
37. Conclusion

Appendix to Chapter IV 283
Frege, Russell, and the Puzzles of Reference
1. Russell's solutions to the puzzles of reference
2. Fregean solutions to the same puzzles
3. Reviewing Fregean assumptions
4. Reviewing Russell's assumptions
5. Building a bridge between both views
6. Conclusion

Chapter V 308
Verificationism Redeemed
1. Origins of semantic verificationism
2. Wittgensteinian verificationism
3. Verifiability rule as a criterial rule
4. Objection 1: the principle is self-refuting
5. Objection 2: a formalist illusion
6. Objection 3: verificational holism
7. Objection 4: existential-universal asymmetry
8. Objection 5: arbitrary indirectness
9. Objection 6: empirical counterexamples
10. Objection 7: formal counterexamples
11. Objection 8: skepticism about rules
12. Defending analyticity
13. Conclusion

Appendix to Chapter V 354
The Only Key to Solving the Humean Problem of Induction
1. Formulating the Humean argument
2. The basic idea
3. Reformulating PF

Chapter VI 366
Sketch of a Unified Theory of Truth
1. The deceptive simplicity of correspondence
2. Analysis of correspondence (1): structural isomorphism
3. Analysis of correspondence (2): categorial match
4. Analysis of correspondence (3): intentionality and causality
5. Exemplifying correspondence
6. Compatibility between verificationism and correspondence
7. Formal definitions of truth
8. Negative truths
9. Self-referentiality
10. Pragmatics of the correspondence relation
11. Retrograde procedures
12. A more complex case
13. General statements
14. Some funny facts
15. Expansion to formal sciences
16. Why can analytic truth be called true?
17. The insufficiency of coherence

18. Coherence as a mediator
19. Roles of empirical coherence
20. Reverend David's case
21. What about the truth of the truth-maker?
22. Objection of the linguistic-cognitive circle
23. Answering the objection of the linguistic-cognitive circle
24. The argument of illusion
25. Answering the argument of illusion
26. The argument of science and its answer
27. Question: How do we warrant the perception of external content?
28. Answer: A definitional criterion of external reality
29. Proving the existence of the external world
30. Skeptical scenarios
31. Verification and intentionality: Husserl
32. Solving two Husserlian problems
33. Truth and factual existence again
34. The rule's structural mirroring of the world
35. Conclusion

Appendix to Chapter VI 458
Discovery of Wine

References 460

It’s not madness that turns the world upside down. It’s conscience.
Bernard Malamud

PREFACE

Niemand weiß noch, wer künftig in jenem Gehäuse wohnen wird und ob am Ende dieser ungeheuren Entwicklung ganz neue Propheten oder eine mächtige Wiedergeburt alter Gedanken und Ideale stehen werden, oder aber – wenn keins von beiden – mechanisierte Versteinerung, mit einer Art von krampfhaftem sich wichtig nehmen verbrämt. Dann allerdings könnte für die "letzten Menschen" dieser Kulturentwicklung das Wort zur Wahrheit werden: "Fachmenschen ohne Geist, Genußmenschen ohne Herz: dies Nichts bildet sich ein, eine nie vorher erreichte Stufe des Menschentums erstiegen zu haben."
[No one knows who will live in this cage in the future, or whether at the end of this tremendous development entirely new prophets will arise, or there will be a great rebirth of old ideas and ideals, or, if neither, mechanized petrification, embellished with a sort of convulsive self-importance. For the 'last man' of this cultural development, it might well be truly said: 'Specialist without spirit, sensualist without heart; this nullity imagines that it has attained a level of humanity never before achieved.']
—*Max Weber*

The aim of philosophy, abstractly formulated, is to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term.
—*Wilfrid Sellars*

Making empty is the result of making small.
—*Malcolm Bull*

Science (mainly applied science) rises, while culture (artistic, religious, philosophical) falls. Whereas culture was once a source of values, today science and technology have made cultural values seem superfluous.

The critical theory of society has offered some explanations for this, drawing on Max Weber's basic idea of the *disenchantment of the world* (*Entzauberung der Welt*). According to him, Western society has undergone a long and seemingly irreversible process of rationalization, in which a scientific-technological society, characterized by increasing bureaucratic rationality, gradually becomes alienated from the values, traditions, and sentiments of older forms of social thinking and acting, without having developed suitable resources to fill the void left behind.

As a result, in a scientifically oriented society, instrumental reason tends to prevail over valuing reason, furthering science and technology at the expense of an adequate substitute for the traditional aesthetic, mystical and humanistic cultural practices, which the available science remains unable to replace. Sociologists have used terms like ‘anomia,’ ‘alienation,’ and ‘nihilism’ to designate the negative individual and social effects of this mismatch between science and humanistic thinking, complaining that our technological world demands forms of cultural alienation to feed itself. Mass culture is a poor attempt to fill the gap; another is scientism.

Given the pressure of modern social forms resulting from rapidly spreading disenchantment, we should not wonder that a kind of philosophy prevails that all too often materially and institutionally simulates the methods and aims of particular scientific fields. In fact, it often emulates the sciences in a manner suggesting the way much of continental philosophy has emulated rhetorical-literary forms, that is, taking over the place of the most proper forms of philosophical argumentation with the effect of losing much of its relation to truth. As a fact, a scientistic attempt to ‘disenchant philosophy’ is incoherent because science in a wide sense must be ‘consensualizable public knowledge’ (John Ziman), opposed in this way by the inevitably non-consensualizable philosophical activity, often turning itself into a mix of pseudo-science and bad philosophy. Hence, a scientistic attempt to disenchant philosophy is, in fact, a thinly veiled attempt of ‘re-enchantment.’ However, it must be a deficient one, insofar as the epistemic place of philosophy in its central domains is by intrinsic necessity deeply ingrained in older forms of a pluralist conjectural argumentative endeavor aiming at comprehensiveness, which cannot be reduced to the domain of a particular science without being severely mutilated.

We can feel this tension in praxis: by taking into account only the discussions of recent years, as science does, one might pretend that the philosophical community is going through the same linear development as science, only to find itself some time later lost in a confusing variety of foreseeable *culs-de-sac*. But an inevitably segmented ‘minute philosophy’ of the ‘last novelty’ made for ‘immediate consumption’ by and for small self-protective cliques of specialists and related scientists no longer seems, as in the tradition, to be an independent conjectural undertaking making balanced use of whatever new scientific knowledge can serve its purposes. More often, it appears often as a busy handmaiden of science suffering from loss of identity and self-esteem; a forcefully particularized pseudo-scientific guesswork, an atomized conjectural endeavor that does not look beyond its own narrow interests. This guesswork scarcely touches the

central philosophical problems inherited from the philosophical tradition or touches them in a way that is unrecognizably deformed by their own reductive-positivist perspective. They seem unprepared to see that in its most central domains philosophy should absorb science instead of being absorbed by science.

In pointing to this, I am far from embracing Manichaeism. I am not claiming that for science to exert great influence on philosophy is inevitably specious and unfruitful. There are many useful limited ways of doing philosophy. Often, particularized philosophy furthers the development of particular sciences or develops into a new field that approaches science, as in the very successful case of speech acts theory. Moreover, there are felicitous cases, like the rapid proliferation of competing theories of consciousness over the last five decades, which serves as a striking example of fruitful philosophical work very closely associated with the development of empirical science that has deepened the field of investigation. And these are only a few cases among many!

Nevertheless, it is important to remember that this same intellectual movement can easily become an ideologically motivated agenda if it tempts the theoretical philosopher to import new knowledge from particular sciences – formal or empirical – in ways that cause him to lose sight of the vast and plural scope of the philosophical landscape. A possible consequence of this is what can be aptly labeled *expansionist scientism*: an effort to reduce some wide domain of philosophy to the scope of investigative strategies and categories derived from a new more or less established particular science. In order to achieve this aim, the particular (formal or empirical) scientific field must be expanded in order to answer questions belonging to some more central domain of philosophy, using a reductionist strategy that underestimates philosophy's encompassing and multifaceted character. An earlier example of expansionist scientism was in my view Pythagoreanism, which unsuccessfully tried to find answers to the problems of life using the newly developed science of numbers. Today's example would be modal logic, which has also generated a fair amount of expansionist scientism. The price one must pay for this may be that persistent, distinctive philosophical difficulties, which cannot be accommodated within the new particularizing model must be minimized if not quietly swept under the carpet.

A chief inconsistency of scientism arises from the fact that while sciences are in various ways all particular, philosophy is most properly 'holistic': As Wittgenstein once wrote, the difficulty of philosophy is that its problems are so interconnected that it is impossible to solve any one philosophical problem without first having solved all the others. Insofar as

his claim is true, it means that a persistent difficulty of the central philosophical problems is that we need a proper grasp of the whole to be able to evaluate and answer them properly. Indeed, this is what can make philosophical understanding so unbearably complex and multifarious. And the lack of this kind of comprehensiveness is what can make fragmented contemporary analytic philosophy often appear like a headless turkey running around aimlessly. Nonetheless, taking account of parts as belonging to a whole, trying to see things *sub specie totius*, is what the great systems of classical philosophy – such as those of Aristotle, Kant, and Hegel – strove to achieve, even if paying a price that we are now better able to appreciate as unavoidably high in terms of misleading and aporetic speculation. Nonetheless, it would be too easy and hasty to conclude that *true comprehensiveness* is no longer a fundamental desideratum of philosophy (Wittgenstein was well aware of this when he called for more '*Übersichtlichkeit*').

There is also an internal reason for the narrowness and fragmentation of much of our present linguistic-analytical philosophy that can be explained as follows. The new Anglo-American philosophy – from W. V-O. Quine to Donald Davidson, and from Saul Kripke to Hilary Putnam and Timothy Williamson – has challenged a great variety of inherited commonsense starting points and challenged them in often undeniably insightful and imaginative ways, although in my view with ultimately unsustainable results. Because of this, a considerable part of theoretical philosophy has increasingly lost touch with its intuitive commonsense grounding in the way things *prima facie* seem to be and for the most part really are.

Take, for instance, the concept of meaning: the word 'meaning' was challenged by Quine as too vague a noise to be reasonably investigated. But an approach is inevitably limited if it, moved by contentious arguments, starts from a kind of positivist-reductionist perspective that denies or ignores commonsense certainties, like the indisputable fact that meanings exist and demand an appropriate explanation. Indeed, using the strategy of skeptically questioning all kinds of deeply ingrained truisms, scientistically oriented philosophers have sawed off the branches they were sitting on. The reason for this is that the result of the adopted strategy couldn't be other than replacing true comprehensiveness with a superficializing positivistic fragmentation of inevitably misleadingly-grounded philosophical concerns. This movement ends by plunging philosophy into what Scott Soames confidently called the 'age of specialization,' while Susan Haack with a healthy touch of pessimism would call it 'a disastrous age of fragmentation.'

Admittedly, this fragmentation can be regarded as dividing to conquer; but it may also be a matter of dividing to subjugate, and what is here to be subjugated is more often the philosophical intellect. Indeed, by focusing too much on the trees, we may lose sight of the philosophical forest and thereby even of where the trees are and how to compare them. Without the well-reasoned assumption of some deep common sense truisms, no proper *descriptive metaphysics*, to use P. F. Strawson's expression, remains possible. And without this, the only path left for originality in philosophy of language, after rigorous training in techniques of argumentation, may turn out to be the use of new formalistic pyrotechnics of unknown value or the production of intellectual artificialities of scarce intelligibility and suspicious depth. This would have the end-effect of blocking paths of inquiry, disarming adequate philosophical analysis and increasing the risk that the whole enterprise will degenerate into a sort of scholastic, fragmented, vacuous intellectual *Glasperlenspiel.*

It may be that practitioners of reductive scientistic philosophy are aware of the problem, but they have found plausible excuses for neglecting to deal with it. Some have suggested that any attempt to do philosophy on a comprehensive level would not suffice to meet the present standards of scholarly adequacy demanded by the academic community. But in saying this they forget that philosophy does not need to be pursued too close on the heels of new advances in the sciences, which are continually producing and handing down new authoritative developments. Philosophy largely remains an autonomous *cultural* enterprise: it is inherently conjectural and dependent on metaphorical elements indispensable to its pursuit of comprehensiveness (Aristotle, calling his first philosophy 'the *searched for* science' was well aware of this). Indeed, most of philosophy remains a relatively free cultural enterprise with a right to controlled speculation, experimentation, and even transgression, though most properly done in the pursuit of truth.

Others have concluded that today it is impossible to develop a truly encompassing theoretical philosophy. For them this kind of philosophy cannot succeed because of the difficulties imposed by the overwhelming amount of information required, putting the task far beyond the cognitive capacity of individual human minds. We might even be – to borrow Colin McGinn's original metaphor – *cognitively closed* to finding decisive solutions for the great traditional problems of philosophy in the sense that we aren't adequately wired to solve them. That is, in our efforts to do ambitious comprehensive philosophy, we are like chimps trying to develop the theory of relativity. Just as they lack sufficient mental capacity to solve the problems of relativistic mechanics, we lack sufficient mental capacity

to develop comprehensive philosophy and will therefore never succeed! Hence, if we wish to make progress, we should shift our efforts to easier tasks...

This last answer seems specious and borders on defeatism. The very ability to initiate the discussion of broadly-inclusive philosophy suggests that we might also be able to accomplish our task. As Wittgenstein once noted, if we are able to pose an appropriate question, it is because we are also in principle able to find its answer. In contrast to human thinkers, one indication that chimps could never develop a theory of relativity is that unlike Einstein they are unable to even pose questions such as what would happen if they could move at the speed of light. Moreover, even if the total amount of scientific knowledge available to us has increased immensely, it may well be that the amount of really essential information needed to answer any given question is sufficiently limited for us to grasp and apply. Very often the science needed to do philosophy can be limited to very general findings. Furthermore, not all philosophical approaches need to be taken into account, since they are often superimposed or displaced. The main difficulty may reside in the circumstances, strategies and authenticity of attempts, in limits imposed on the context of discovery, rather than in the sheer impossibility of progress. In any case, it is a fact that in the so-called philosophy of linguistic analysis true comprehensiveness has almost disappeared in the recent years. However, my guess is that the main reason isn't impossibility in principle, but rather *the loss of a suitable cultural soil in which a more comprehensive philosophy could flourish*.

In this book, I begin by arguing that more fruitful soil can be found if we start with a better reasoned and more affirmative appreciation of commonsense truisms, combined with a more pluralistic approach, always prepared to incorporate the relevant – formal and empirical – results of science. Perhaps it is precisely against the uncomfortable return of a broader pluralistic approach that much of the mainstream of our present philosophy of language secretly struggles. Awareness of this can be obscured by some sort of dense, nearly scholastic scientistic atmosphere, so thick that seasoned practitioners barely notice it surrounding them. The intellectual climate sometimes recalls the Middle Ages, when philosophical investigation was allowed, providing it left unchallenged established religious dogmas. I even entertain the suspicion that in some quarters the attempt to advance any plausible comprehensive philosophy of language against the institutional power of reductive scientism runs the risk of being ideologically discouraged as a project and silenced as a fact.

Ernst Tugendhat, who (together with Jürgen Habermas) attempted with considerable success to develop comprehensive philosophy in the

seventies, seems to have hoisted the white flag by admitting that the heyday of philosophy is past. The problem is in my view aggravated because we live in a time of widespread indifference concerning high culture, as I pointed out at the beginning – a time heavily influenced by a steady, almost exponential development of science and technology that forcefully minimizes the role of valuing reason. Though quite indispensable from the viewpoint of instrumental reason, our scientifically biased age tends to impose a compartmentalized form of alienation on philosophical research that works against more broadly oriented attempts to understand reality.

In the present book, I insist on swimming against the tide. My main task here – a risky one – is to establish the foundations of a more comprehensive philosophy of meaning and reference, while arguing against some main reductionist-scientistic approaches that are blocking the most promising paths of inquiry. Hence, it is an attempt to restore its deserved integrity to the analytic philosophy of language, without offending either common sense or science; an effort to give a balanced, systematic and sufficiently plausible overview of meaning and the mechanisms of reference, using bridges laboriously constructed between certain summits of philosophical thought. In this way, I hope to realize something of the old philosophical ambition of a comprehensive synthesis, insofar as this still sounds like a reasonable undertaking.

Paris, 2017

ACKNOWLEDGMENTS

Before any acknowledgments, I must emphasize Wittgenstein's major influence on my philosophical outlook. His exceedingly suggestive and multifarious approach is more far-reaching than unprepared readers could possibly grasp, and the originality of his philosophical mind is indebted to his freedom from the burdens of the academic factory. Frege's unprecedented work was equally indispensable in structuring the subject under investigation. I must also emphasize as well the strong influence on my thinking of two living philosophers: Ernst Tugendhat and John Searle.

Regarding proper acknowledgments, first I would like to thank Anna-Sofia Maurin and her talented students from the University of Göteborg in Sweden for critically discussing with me some main issues of a later draft of this book and asking challenging questions. I am also especially grateful to François Recanati's generous invitation to spend time at the Institute Jean Nicod in Paris; his teaching of pragmatics was for me very instructive. I am also indebted to the organizers of the 40th *Wittgenstein Symposium* in Kirchberg am Wechsel for the opportunity of presenting some views on philosophy of perception that I tried to better develop in this book.

Earlier partial drafts I discussed years ago with Wolfgang Spohn at the University of Konstanz in Germany and presented to João Branquinho and his colleagues at the University of Lisbon. I am grateful to each of them for their helpful suggestions. I am also grateful to Guido Imaguire, Dirk Greimann, Ethel Rocha and faculty members at the Department of Philosophy of the IFCS in Rio de Janeiro for their remarks. Moreover, I would like to express my gratitude to several persons who in different ways have influenced this work: Richard Swinburne, Allan Sidelle, Cheryl Misak, David Rosenthal, Susan Haack, Manuel García-Carpintero, Marco Ruffino, André Leclerc and Nelson Gomes. Cinara Nahra, as head of my department at the UFRN, has greatly supported and encouraged my work, and my colleague Daniel Durante has made valuable suggestions on numerous occasions. I warmly thank my former professors Raul Landin and Guido de Almeida for encouraging my work on this project. And I would also like to thank Dr. James Stuart Brice for carefully and insightfully proofreading my text. It goes without saying that I am solely responsible for the views defended in this book.

CHAPTER I

INTRODUCTION

> Logic, I should maintain, must no more admit a unicorn than zoology can; for logic is concerned with the real world just as truly as zoology, though with its more abstract and general features.
> —*Bertrand Russell*

> A philosophical tradition which suffers from the vice of *horror mundi* in an endemic way is condemned to futility.
> —*Kevin Mulligan, Peter Simons, Barry Smith*

The old orthodoxy of the philosophy of language that prevailed during the first half of the twentieth century was marked by an insistence on the centrality of meaning, an eroded semantic principle of verifiability, naive correspondentialism, an elementary distinction between analytic and synthetic, crude descriptivist-internalist theories of proper names and general terms, a monolithic dichotomy between the *necessary a priori* and the *contingent a posteriori*… Could it nevertheless come closer to the truth than the now dominant causal-externalist orthodoxy?

This book was written in the conviction that this question should be answered affirmatively. I am convinced that the philosophy of language of the first half of the twentieth century that formed the bulk of the old orthodoxy was often more virtuous, more comprehensive, more profound and closer to the truth than the approaches of the new orthodoxy, and that its rough-hewn insights were often more powerful, particularly in the works of philosophers like Wittgenstein, Frege, Russell and even Husserl. My conjecture is that the reason lies in the socio-cultural background. Even if also motivated by a desire to approach the authentic consensual truth only possible for science, philosophy in itself has its own epistemic place as a cultural conjectural endeavor, unavoidably harboring metaphorical components which can be approached to those of the fine arts and comprehensive aims approachable to those of religion, even if it is in itself independent of both (Costa 2002). In its best, the first half of the twentieth century preserved these traits. One reason might be that this was still a very elitist and hierarchical intellectual world, while our present academic

world is much more level by a scientifically oriented pragmatic society, which does not make it the best place for philosophy as an effort to reach surveillability. A more important reason is that great culture is the result of a great conflict. And the period between the end of the nineteenth century and the Second World War was a time of increasing social turmoil with tragic dimensions. This conflict cast doubt on all established cultural values, creating the right atmosphere to the emergence of intellectuals and artists disposed to develop sweepingly original innovations. This could be felt not only in philosophy and the arts, but also in fields reserved for particular sciences.

Philosophy of language since the Second World War has been much more a form of strongly established academic 'normal philosophy,' to borrow Thomas Kuhn's term. On the one hand, it was a continuation of the old orthodoxy, represented in the writings of philosophers like John Austin, P. F. Strawson, Michael Dummett, John Searle, Ernst Tugendhat, Jürgen Habermas… whose side I usually take. On the other hand, we have seen the emergence of what is called the new orthodoxy, founded by Saul Kripke and Keith Donnellan in the early seventies and later elaborated by Hilary Putnam, David Kaplan, and many others. In opposition to the old orthodoxy, this approach emphasizes externalism about meaning, causalism, and anti-cognitivism. This new orthodoxy has become the contemporary mainstream position in philosophy of language.

I do not deny the philosophical relevance of this new orthodoxy. Nor do I reject its originality and dialectical force. Perhaps I am more indebted to it than I wish to admit. Nevertheless, it has already long since lost much of its creative impetus, and it now has transformed itself into a kind of scholastic discussion among specialists. Moreover, the value of the new orthodoxy in philosophy of language is in my judgment predominantly negative, since most of its conclusions fall short of the truth. This means that the significance of its ideas consists mostly in their being dialectically relevant challenges, which, I believe, could be adequately answered by an improved reformulation of old, primarily descriptivist-internalist-cognitivist views of meaning and its connection with reference that are to some extent developed in the present book. Indeed, I intend to show that the views of the old orthodoxy could be reformulated in much more sophisticated ways, not only answering the challenges of the new orthodoxy, but also suggesting solutions to problems that the contemporary philosophy of language hasn't addressed as well as it should.

My approach to the topics considered here consists in gradually developing and defending a primarily internalist, cognitivist and neodescriptivist analysis of the nature of the cognitive meanings of our expressions and

their inherent mechanisms of reference. But this approach will be indirect since the analysis will be supported by a critical examination of some central views of traditional analytic philosophy, particularly those of Wittgenstein and Frege. Furthermore, such explanations will be supplemented by a renewed reading and defense of the idea that existence is a higher-order property, a detailed revaluation of the verificationist explanation of cognitive meaning, and a reassessment of the correspondence theory of truth, which I see as complementary to the here developed form of verificationism, involving coherence and dependent on a correct treatment of the epistemic problem of perception.

The obvious assumption that makes my project *prima facie* plausible is the idea that language is a system of rules, some of which should be the most proper sources of meaning. Following Ernst Tugendhat, I assume that the most central meaning-rules are those responsible for what Aristotle called *apophantic speech*: the representational discourse, whose meaning-rules I call *semantic-cognitive rules*. Indeed, it seems at first highly plausible to think that the cognitive meaning (i.e., informative content and not mere linguistic meaning) of our representational language cannot be given by anything other than semantic-cognitive rules or associations of such rules. Our knowledge of these typically conventional rules is – as will be shown – usually tacit, implicit, non-reflexive. That is, we are able to use them correctly in a cognitive way, though we find almost unsurmountable difficulties when trying to analyze them in a linguistically explicit way, particularly when they belong to philosophically relevant concepts.

My ultimate aim should be to investigate the structure of semantic-cognitive rules by examining our basic referential expressions – singular terms, general terms and also declarative sentences – in order to furnish an appropriate explanation of their reference mechanisms. In the present book, I do this only partially, often in the appendices, summarizing ideas already presented in my last book (2014, Chs. 2 to 4), aware that they still require development. I proceed in this way because in the main text of the present book my main concern is rather to justify and clarify my own assumptions on the philosophy of meaning and reference.

1. Ernst Tugendhat's analysis of singular predicative statements

In developing these views, I soon realized that my main goal could be seen as essentially a way to revive a program already speculatively developed by Ernst Tugendhat in his classical work *Traditional and Analytical*

Philosophy: Lectures on the Philosophy of Language.[1] This book, first published in 1976, can be considered the swansong of the old orthodoxy, defending a non-externalist and basically non-properly-causalist program that was gradually forgotten during the next decades under the ever-growing influence of the new causal-externalist orthodoxy. Tugendhat's strategy in developing this program can be understood in its core as a semantic analysis of the fundamental singular predicative statement. This statement is not only epistemically fundamental, it is also the indispensable basis for building our first-order truth-functional language. In summary, given a statement of the form *Fa*, he suggested that:

1) The meaning of the singular term *a* should be its *identification rule* (*Identifikationsregel*),
2) the meaning of the general term *F* should be its *application rule* (*Verwendungsregel*), which I also call a *characterization* or (preferably) an *ascription* rule,
3) the meaning of the complete singular predicative statement *Fa* should be its *verifiability rule* (*Verifikationsregel*), which results from the collaborative application of the first two rules.

(*Cf.* Tugendhat & Wolf 1983: 235-6; Tugendhat 1976: 259, 484, 487-8).

In this case, the verifiability rule is obtained by the sequential application of the first two rules in such a way that the identification rule of the singular term must be applied first, in order to then apply the general term's ascription rule. Thus, for instance, Yuri Gagarin, the first man to orbit the Earth from above its atmosphere, gazed out of his space capsule and exclaimed: 'The Earth is blue!' In order to make this a true statement, he should first have identified the Earth by applying the identification rule of the proper name 'Earth.' Then, based on the result of this application, he would have been able to apply the ascription rule of the predicative expression '…is blue.' In this form of combined application, these two rules work as a kind of verifiability rule for the statement 'The Earth is blue.' That is: if these rules can be conjunctively applied, then the statement is true, otherwise, it is false. Tugendhat saw this not only as a form of verificationism, but also as a kind of correspondence theory of truth – a conclusion that I find correct, although rejected by some of his readers.[2]

---

[1] Original German title: *Vorlesungen zur Einführung in die sprachanalytische Philosophie*.

[2] An antecedent of this is J. L. Austin's correspondence view, according to which an indexical statement (e.g., 'This rose is red') is said to be true when the historical

In order to test Tugendhat's view, we can critically ask if it is not possible that we really first apply the ascription rule of a predicative expression. For example, suppose that one night you see something burning at a distance without knowing what is on fire. Only after approaching it do you see that it is an old, abandoned factory. It may seem that in this example you first applied the ascription rule and later the identification rule. However, in suggesting this you forget that to see the fire one must first direct one's eyes at a certain spatiotemporal spot, thereby localizing the individualized place where something is on fire. Hence, a primitive identification rule for a place at a certain time needed to be first generated and applied.

That is, initially the statement will not be: 'That old building is on fire,' but simply 'Over there… is fire.' Later on, when you are closer to the building, you can make a more precise statement. Thus, in this same way, while looking out of his space capsule's porthole, Gagarin could think, 'Out there below the porthole it is blue,' before saying 'The Earth is blue.' But even in this case, the ascription rule cannot be applied without the earlier application of some identification rule, even if it is one that is only able to identify a vague spatiotemporal region from the already identified porthole. To expand on the objection, one could consider a statement like 'It is all white fog.' Notwithstanding, even here, 'It is all…' expresses an identification rule (of my whole visual field covering the place where I am right now) for the singular term, while '…white fog' expresses the ascription rule that can afterward be applied to the whole place where I am. Even if there is no real property, as when I state 'It is all darkness,' what I mean can be translated into the true statement 'Here and now there is no light.' And from this statement, it is clear that I first apply the indexical identification rule for the here and now and afterward see the inapplicability of the ascription rule for lightness expressed by the negation '…there is no light' corresponding to the predicative expression '…is all darkness.'

Tugendhat reached his conclusions through purely speculative considerations, without analyzing the structure of these rules and without answering the many obvious external criticisms of the program, like the numerous well-known objections already made against verificationism.

---

fact correlated with its demonstrative convention (here represented by the demonstrative 'this') is of the type established by the sentence's descriptive convention (the red rose type) (Austin 1950: 122). This was a first approximation of conventionalist strategies later employed by Dummett in his interpretation of Frege (*Cf.* 1981: 194, 229) and still later more cogently explored by Tugendhat under some Husserlian influence.

But what is extraordinary is that he was arguably right, since the present book will make it hard to contest his main views.[3]

2. The virtue of comprehensiveness

Our methodological strategies will also be different from those used in the more formalistically oriented approaches criticized in this book, insofar as they follow a positivist-scientistic kind of ideal language philosophy that often hypostasizes form in ways that lead them to ignore or distort empirical truisms. By contrast, I am more influenced by what could be broadly called the *natural language* tradition, thus being inclined to assign a fair amount of heuristic value to common sense and critical examination of the natural language intuitions, often seeking support in a more careful examination of concrete examples of how linguistic expressions are effectively employed in adequately chosen conversational contexts.[4] Consequently, my approach is primarily oriented by the communicative and social roles of language, which are regarded as the fundamental units

[3] Tugendhat's thesis crosses over peculiarities of linguistic interaction. Consider a conversational implicature: – 'Do you know how to cook?' – 'I am French,' which implicates the statement 'I know how to cook.' (Recanati 2004: 5) Obviously, this does not effect Tugendhat's thesis, for the proper and implied meanings posed by the statement 'I am French' would then be established by means of verifiability rules.

[4] The ideal language tradition (steered by the logical analysis of language) and the natural language tradition (steered by the real work of natural language) represent opposed (though arguably also complementary) views. The first was founded by Frege, Russell and the early Wittgenstein. It was also later strongly associated with philosophers of logical positivism, particularly Rudolf Carnap. With the rise of Nazism in Europe, most philosophers associated with logical positivism fled to the USA, where they strongly influenced American analytic philosophy. The philosophies of W. V-O. Quine, Donald Davidson, and later Saul Kripke, Hilary Putnam and David Kaplan, along with the present mainstream philosophy of language, with its metaphysics of reference, are in indirect ways later American developments of ideal language philosophy. What I prefer to call the natural language tradition was represented after the Second World War in Oxford by the sometimes dogmatically restrictive 'ordinary language philosophy.' Its main theorists were J. L. Austin, Gilbert Ryle, and P. F. Strawson, although it had an antecedent in the less restrictive natural language philosophy of the later Wittgenstein and, still earlier, in G. E. Moore's commonsense approach. Natural language philosophy also affected American philosophy through relatively isolated figures like Paul Grice and John Searle, whose academic influence has foreseeably not been as great... For the initial historical background, see J. O. Urmson (1956).

of analysis. It must be so because I assume that the most properly philosophical approach should be as comprehensive as possible and that an all-inclusive understanding of language and meaning must fairly contemplate its unavoidable involvement in overall societal life.

Finally, my approach is systematic, which means that coherence belongs to it heuristically. The chapters of this book are so interconnected that the plausibility of each is usually better supported when regarded in its relation to arguments developed in the preceding chapters and their often critical appendices. Even if complementary, these appendices (particularly the Appendix of the present introduction) are sometimes an indispensable counterpoint to the chapters, aiming to better justify the expressed views, if not to add something relevant to them.

The whole inquiry strives in the direction of comprehensiveness, aiming to reintegrate theoretical philosophy under the recognition that there is no philosophical question completely independent of all the others.[5] In this way, it shows itself to be an attempt to analyze linguistically approximated concepts like meaning, reference, existence, and truth, insofar as they are internally associated with one another and, unavoidably, with a cluster of some main metaphysical and epistemological framework concepts constitutive of our understanding of the world.

[5] After his broad exposition of contemporary philosophy, K. A. Appiah concluded: 'The subject is not a collection of separate problems that can be addressed independently. Issues in epistemology and the philosophy of language reappear in the discussions of philosophy of mind, morals, politics, law, science, and religion… What is the root of the philosophical style is a desire to give a *general* and *systematic* account of our thought and experience, one that is developed critically, in the light of evidence and arguments.' (2003: 377-378) Because of this, the hardest task for those committed to comprehensive coherence is to reach a position that enables the evaluation of the slightest associations among issues belonging to the most diverse domains of our conceptual network (*Cf.* Kenny 1993: 9).

APPENDIX TO CHAPTER I

HOW DO PROPER NAMES REALLY WORK? (CUTTING THE GORDIAN KNOT)

> Once fashion comes in, objectivity goes.
> —*D. M. Armstrong*

As Wittgenstein once said, our aim in teaching philosophy should not be to give people the food they enjoy, but rather to offer them new and different food in order to improve their tastes. This is my intention here. I am firmly convinced that I have a much more elucidative explanation for the mechanisms of reference that characterize proper names, but the really difficult task seems to be that of convincing others. This difficulty is even greater because I am swimming against the present mainstream – in this case, the externalist-causalist and basically anti-cognitivist views regarding the meaning and reference of proper names.

There is a further reason why the neodescriptivist theory of proper names that I intend to summarize here is particularly hard to accept. This is because the question of how proper names refer has always been the touchstone for theories of reference. More than forty years ago, when Saul Kripke, Keith Donnellan, and others rejected descriptivism for proper names, they also opened the door for externalist, causalist, and potentially non-cognitivist views concerning the reference of indexicals, natural kind terms, and statements. Now, if I achieve my goal, which is to re-establish descriptivism concerning proper names in a considerably more developed and refined way, the doors will again be open to re-establishing descriptivist-internalist-cognitivist views about other terms and language in general. This means that we will once again have to survey the whole topography of our philosophy of language. However, since the new orthodoxy is already well-entrenched – it has led a good life for the past forty years – and a myriad of good and bad arguments have been developed in its favor, the challenge is naturally huge. If I limited myself to answering just the most relevant arguments, I would still need to write an entire book to make a persuasive case for a neodescriptivist approach to proper names. But when I consider the potential disorder that advanced

neo-descriptivism could cause in all these 'well-established' views about reference, even a thousand-page book defending the descriptivist-internalist-cognitivist understanding of terms and answering all the relevant arguments still seems insufficient. And the reason is clear: most specialists are now working within the externalist-causalist paradigm, and many do not wish to be convinced. Taking into consideration that I am not writing for readers with unshakable theoretical commitments, in what follows I dare to offer a summarized version of my own view on proper names.[1]

1. A meta-descriptive rule for proper names

According to descriptivism, proper names are *abbreviations* of definite descriptions. The most explicit formulation of descriptivism for proper names – the *bundle theory* as presented in the work of John Searle – states that a proper name abbreviates a bundle of definite and even indefinite descriptions that constitute its whole content (1958; 1967). This means that definite descriptions have no function other than to be *carriers of information* that can be more or less helpful for the identification of their bearers. As Susan Haack wrote, summarizing Searle's view:

> The different senses we can give a proper name that we use result from our having in mind some not previously determined sub-bundle from a whole bundle of co-referential descriptions. (*Cf.* Haack 1978: 58)

Thus, as Frege already saw, one speaker can use the name 'Aristotle' to mean 'the greatest disciple of Plato and the tutor of Alexander,' while another can use it to mean 'the tutor of Alexander who was born in Stagira' (1892: 29). And in the usual case, both speakers can know they are referring to the same person, insofar as they know that they share at least one description (Frege 1918: 65).

In my view, the problem with this formulation of bundle theory is not that it is wrong, since in one way or another most objections to it can be answered (*Cf.* Searle 1983, Ch. 9). The problem is that this theory is too vague, for this reason lacking explanatory power. The descriptions belonging to the bundles are treated as if they were completely *disordered*. How important this is becomes apparent when we remember that the descriptions belonging to these bundles can be seen as what Wittgenstein called 'expressions of rules' (*Regelausdrücke*): *description-rules* that

[1] This appendix is to some extent a summarized version of a more detailed text entitled 'Outline of a Theory of Proper Names' (Costa 2014, Ch. 2).

could possibly aid us to identify the bearer of a proper name. Usually, there are numerous descriptions that could be associated with any proper name, many of them obviously irrelevant. Unfortunately, bundle theory has no method for deciding which description-rules belonging to a bundle have more relevance for the identification of a name's bearer. It thus appears that the lack of such a method is the most serious flaw in traditional bundle theory.

Accordingly, my working hypothesis is that speakers of our language implicitly appeal to some kind of general *meta-descriptive rule* when using a proper name. This rule should tell us the conditions under which satisfaction of descriptions belonging to a bundle of descriptions abbreviated by a proper name makes this name applicable to its bearer. Thus, I intend to show that such an additional rule can be discovered as part of the pre-existing tools of our natural language and that its full explanation would greatly enhance the bundle theory of proper names.

The first move in this direction should be to find the most relevant descriptions. My proposal is inspired by J. L. Austin's method of quasi-lexicographical examination of ordinary language as a philosophical starting point. He recommended beginning with the *Oxford Dictionary*. Since dictionaries aren't the best places to find the meanings of proper names, I suggest first looking at encyclopedia entries for proper names. By doing this we can clearly distinguish two general kinds of description-rules that can help identify the bearer of a proper name. I call them *auxiliary* and *fundamental* descriptions. Fundamental descriptions are usually placed at the start of encyclopedia articles.

I begin with less relevant auxiliary descriptions. These can be characterized as ones only *accidentally* associated with proper names. Regarding the name 'Aristotle,' typical examples are (i) *metaphorical descriptions* like Dante's 'the master of those who know.' Other examples of auxiliary descriptions are 'the greatest disciple of Plato,' 'the tutor of Alexander,' 'the founder of the Lyceum' and 'the man called "Aristotle."' These are what we may call (ii) *accidental, but well-known descriptions*. There are also (iii) *accidental and little-known descriptions* associated with the name 'Aristotle,' such as 'the lover of Herphyllis' and 'the grandson of Achaeon.' Finally, there are (iv) contextually dependent *adventitious descriptions*, like 'the philosopher mentioned by the professor in the last class,' or 'the blonde woman who spoke with us at the party.' An adventitious description is often very transitory, as it is closely associated with an event that in most cases will soon be forgotten.

Descriptivist philosophers like Frege and Wittgenstein have often used auxiliary descriptions to exemplify parts of a bundle. However, this can be

very misleading, since ultimately they are of negligible semantic relevance. An indication of this secondary role is found in encyclopedias and biographies. Biographies and autobiographies offer a wide range of auxiliary descriptions, mostly irrelevant for identification purposes. Encyclopedias seldom begin articles with auxiliary descriptions. Instead, they begin with what I call *fundamental descriptions*: non-accidental descriptions that usually tell us the 'when' the 'where' and the 'why' of proper-name bearers. Following this path, I define fundamental descriptions as being of the following two types:

(A) *Localizing description-rule:* a description that localizes an object in space and time, often singling out its spatiotemporal career.
(B) *Characterizing description-rule:* a description that indicates what we regard as the most important properties related to the object, exposing our reasons for applying the proper name to it.

Indeed, as a rule, encyclopedias first state a spatiotemporal location and then the main reasons we use a proper name; only after that do they give a more detailed exposition containing most of the auxiliary descriptions. One example is the reference to Aristotle in my short *Penguin Dictionary of Philosophy*, which begins:

> *Aristotle* (384-322 BC) born in Stagira, north of Greece, he produced the major philosophical system of Antiquity…

What we first see here are in synoptic form the localizing and characterizing descriptions.

Having discovered the two most fundamental kinds of description-rules, and after considering several different alternatives that I cannot go into here, I offer the following meta-descriptive rule to establish conditions of application for most if not all proper names. Using the term *world-circumstance* to designate any possible world, including the actual world, not only as we think it is, but also as it could be discovered to be,[2] I can present the meta-descriptive rule as follows:

[2] With the label 'world-circumstance' I wish to make explicit the possibility of discovering errors in our information about the actual world, as in the improbable cases in which we discover that Aristotle was in fact not called 'Aristotle' or in which we discover he was in fact not born in Stagira, etc.

MD-rule for the application of proper names:
In any world-circumstance where a proper name called 'N' has a bearer, this bearer must:
(i) belong to some most proximally relevant class C, so that it
(ii) sufficiently and
(iii) more than any other referent satisfies
(iv) the conditions set by at least (A) its localizing description-rules and/or (B) its characterizing description-rules.
(v) We may add to this, as helpful indicative elements, a variety of auxiliary descriptions.

I illustrate my proposal with the name 'Aristotle.' The (i) most proximally relevant class C to which Aristotle belongs is that of human beings (C serves for practical aims to narrow the scope of referents to be considered, e.g., it excludes celestial bodies or computers). To be more precise, C must be the nearest most relevant class that does not merge with the characterizing description. This is why for the name Aristotle C must be the condition of being *a human being* and not of being *a philosopher*. The condition of type (A) for Aristotle can be summarized by the definite description 'the person born in Stagira in 384 BC, son of the court physician Nicomachus, who spent the most productive part of his life in Athens, visited Lesbos and was exiled to Chalcis, where he died in 322 BC…' The condition of type (B) for Aristotle can be summarized in the definite description 'the philosopher who developed the relevant ideas of the Aristotelian opus…' (That these two conditions are the most basic is supported by major encyclopedias).

Now, by applying the general meta-descriptive rule to the bundle of descriptions abbreviated by the name 'Aristotle,' I finally arrive at what I call its specific *identification rule*, the IR-Aristotle.[3] Summarizing the descriptions, here is the identification rule for Aristotle:

IR-Aristotle: In any world-circumstance where there is a bearer of the proper name 'Aristotle,'[4] this bearer must be: (i) the human being who

[3] We can also read the MD-rule simply as the *form* that any IR-rule for a proper name needs in order to establish its referential condition. (*Cf.* Ch. IV, sec. 15)
[4] I do not identify the name with its symbolic form, but with the identification rule combined with *some* symbolic form. Hence, I place the proper name in quotation marks to indicate that it must be possible to be misleading about the true symbolic form of a proper name. Imagine a possible world where only one philosopher satisfies the fundamental conditions for being our Aristotle, but who is called 'Pitacus.' We would after all still identify him with our Aristotle! Indeed, even in

(ii) sufficiently and (iii) more than any other satisfies (iv) the condition of having been born in Stagira in 384 BC, son of the court physician Nicomachus, spent the major part of his life in Athens and died in Chalcis in 322 BC and/or the condition of being the philosopher who developed the main ideas of the Aristotelian opus. To this, we may add (v) possibly orienting auxiliary descriptions like 'the greatest disciple of Plato,' 'the founder of the Lyceum,' etc.

We can do the same with bundles of descriptions associated with any other well-known proper name, such as 'Paris,' 'Leaning Tower of Pisa,' 'Amazon River,' 'Uranus,' 'Sweden,' and, of course, also with the proper names of historically anonymous persons like most of us, though in the last case in a much more dispersive way that I cannot consider here.

This is my basic idea to explain the mechanism of reference of proper names. Why and how this idea is explanatorily powerful is what I intend to show in the next sections.

2. Identification rules at work

The application of the meta-descriptive rule to the name 'Aristotle' in order to obtain its identification rule enables us to give a straightforward answer to Kripke's modal objection, according to which descriptivism is false, since no description is guaranteed to apply to any existing bearer of a proper name. As he puts it, there could be possible worlds where Aristotle lived 500 years later or where he died as a child and never wrote anything about philosophy (Kripke, 1980: 62 f.). However, these possibilities are no threat to the rule stated above, since this rule is based on an *inclusive disjunction*. Aristotle would still be Aristotle if he had lived 500 years later in Rome, insofar as he sufficiently satisfied the characterizing description related to his work, for instance, by writing major parts of the Aristotelian opus. He could also have died as child, as long as he sufficiently satisfied the localizing description, for instance, birth in Stagira in 384 BC as the son of the court physician Nicomachus.

---

our real world, we cannot completely exclude the possibility that our Aristotle was in fact called Pitacus… This would be true even if Aristotle had no name except 'Someone.' What individuates a proper name is the identification rule we associate with it. The description 'the man named "Aristotle,"' popular in the so-called metalinguistic theory of proper names, is only a well-known (accidental) auxiliary rule.

Since our identification rule for Aristotle demands only sufficient satisfaction of an inclusive disjunction of the two fundamental descriptions (which purposely does not make any precise specification of degree), we can easily regard the two above considered possibilities as satisfying the identification rule.

Indeed, there are even proper names that typically satisfy only one description-rule of the disjunction. These are 'one-foot' (i.e., having only one of the usual two descriptions) proper names like 'universe' (which contains all that exists and thus can have no localizing description) and 'Z' understood as the arbitrary name for the center of a circle (and without any relevant characterizing description). There are even proper names able to satisfy one term of a disjunction much more than the other, as in the case of a numbered planet of the solar system. Venus, for instance, must satisfy the localizing description-rule requiring that it must be the second planet from the sun, orbiting between Mercury and the Earth for a sufficient period of time… Even if for some reason it has changed its orbit or has lost part its atmosphere or part of its mass, it remains Venus. One could say that the essential element of its characterizing description is already built into its localizing description, namely, the condition that it is a planet.

The only inconceivable alternative is that neither the localizing nor the characterizing description-rule is satisfied to a minimal degree. Such a case was fancifully suggested by John Searle in the following example:

> If a classical scholar claimed to have discovered that Aristotle was no philosopher and wrote none of the works attributed to him, but was in fact an obscure Venetian fishmonger of the late Renaissance, then the 'discovery' would become a bad joke. (1967: 490)

Clearly, no sane person would agree with Searle's classical scholar. Such an illiterate man could not be our Aristotle; the obvious reason is that the fishmonger does not at all satisfy either the localizing or characterizing descriptions.

Two other important elements of the MD-rule for proper names need some clarification. These are what I call the condition (ii) of *sufficiency*, namely, the satisfaction of the inclusive disjunction to a sufficient degree, and the condition (iii) of *predominance*, namely, that it satisfies the inclusive disjunction more than any other competitor.

First, take the condition of *sufficiency*. We can imagine a possible world where there was one Aristotle who was born in 384 BC in the court of Stagira… but died when he was seventeen because his ship sank in a storm while he was crossing the Aegean to study in Athens under Plato. He may have been only an Aristotle *in potentia*, but we would still believe

he was *our* Aristotle! The reason is that the identification rule is satisfied insofar as the localizing conditions are at least sufficiently satisfied (if only partly). It is irrelevant that the other requirement of the disjunction is not satisfied at all. The opposite case would be that of a possible world where the only Aristotle was born 500 years later than ours, lived in Rome and wrote only the *Metaphysics* and the *Nicomachean Ethics.* We would still tend to identify him as our Aristotle. We can of course imagine a possible world where the only Aristotle was born in Stagira in 384 BC, wrote Aristotle's now lost earlier dialogues and the *Organon* and died prematurely before writing the *Metaphysics* and other major works. In this case, both rules are only partially but at least sufficiently satisfied, since we can still identify him as our Aristotle.

The second condition, *predominance*, reveals its purpose when we imagine that the court physician Nicomachus fathered twin sons in Stagira in 384 BC, naming both 'Aristotle.' The first Aristotle moved to Athens when he was seventeen to study philosophy with Plato and later wrote the entire Aristotelian opus. The second Aristotle had less luck... He became a physician like his father and accompanied Alexander on his military campaigns, but succumbed to hunger and thirst in the desert while returning from the East. Who would be our Aristotle? The first one, of course. The reason is that much more than his brother he satisfies the fundamental conditions of the identification rule for Aristotle. The condition of predominance serves to exclude the possibility that more than one object satisfies the identification rule.

If there is more than one object that satisfies the identification rule to the same degree, even if in different ways, our criterial tool for the application of a proper name will fail. Imagine, for instance, a possible world in which it is very common (and normal) for people to have two heads. Suppose that there was a child with two identical heads who was born in the court of Stagira in 384 BC, son of Nicomachus. The two heads developed into separate 'persons' and both were called 'Aristotle.' Since they shared the same body, these two persons inevitably lived very closely linked lives, jointly writing the entire Aristotelian opus. It would be almost pointless to ask which was *our* Aristotle, since by definition proper names apply to only one bearer, and the two satisfy the identification rule in equal measure (however, we could still adopt the strategy of spatially distinguishing the Aristotle 'on the right side' from the Aristotle 'on the left side'…).

The inclusion of the condition of predominance already has the advantage of explaining why it is intuitive for us that a Twin-Earth Aristotle (who was qualitatively identical with our Aristotle and lived on

identical planets…) is not the 'true' Aristotle. This approach works better than Searle's attempted explanation (1983: 254-5) because our earth's Aristotle is the person who satisfies the condition of predominance. Both Aristotles satisfy the characterizing description-rule (they both wrote the *Corpus Aristotelicum*), and because the spatial context is similar, the Twin-Earth Aristotle also appears to satisfy the localizing rule. However, beyond this only the Earth-Aristotle really satisfies the localizing description-rule, since he lived in the Greece of *our* Earth, and not in the far distant Twin-Earth Greece. Because both Earths belong to one and the same space, the localizing description-rule refers to a spatial location on our Earth and not to its counterpart on the distant Twin-Earth, notwithstanding the confusingly similar local surroundings. (Even if there could be two incommensurably different spaces, the conclusion would remain the same, since our Aristotle would belong to the first and not to the second space.)

The introduction of the so understood identification rule allows us to solve the famous paradox of Theseus's ship.[5] Suppose Theseus had a ship named 'Calibdus' that over time showed signs of wear. He gradually replaced its planks with new ones, until in the end there was not one original plank left. All the worn-out planks had been stored, and someone decided to repair them and build a 'new' ship, identical with the original one. Which ship is now the Calibdus? (If you think it must be the first one, you need only speed up the substitution of the planks: if the whole substitution requires just one month, you would tend to say that the second ship is the Calibdus, and if it takes just one day, you will be sure of this.)

This imagined situation seems paradoxical, insofar as it leads us to grasp a possible conflict in the application of the two fundamental description-rules. The first ship better satisfies the localizing description concerning its date of launching and spatiotemporal career; it also satisfies enough of the characterizing description concerning its structural and functional properties, though not its material constitution. The second ship satisfies the characterizing description better since besides its structural and functional properties it has all the original material parts. Both satisfy conditions of sufficiency for Calibdus, but since the paradox invites us to consider a situation in which neither of them satisfies the condition of predominance, we feel that there are cases in which the identification rule isn't applicable any longer, cases in which Calibdus no longer exists in the sense demanded by a singular term.

[5] See Wiggins 2001: 93 f.

One could, finally, ask if auxiliary descriptions play some role regarding predominance. The answer seems to be divided. The answer is 'no' in cases where auxiliary descriptions are irrelevant, like 'the man called by Dante the master of those who know' or 'the grandson of Achaeon.' But in the case of relevant auxiliary descriptions like 'the greatest disciple of Plato' they do seem to matter. And there are limitrophe descriptions like 'the son of Nicomachus, the court physician of Philip of Macedon' that make a clear difference (in the case of several persons named 'Aristotle' born in 384 BC in Stagira we would choose the person who satisfies the above description, unless this person does not satisfy the identification rule more than any other). The border between fundamental and auxiliary descriptions is vague, particularly for the class of proper names of historically irrelevant persons, which includes the great majority of people.

Finally, the insignificance of most auxiliary descriptions comes to the fore when we consider the case of someone who satisfies them but does not satisfy the fundamental conditions. Consider, for instance, the famous Greek shipping magnate Aristotle Onassis (1906-1975). He could not be our Aristotle, because he satisfies none of the fundamental descriptions. But suppose he satisfied some auxiliary descriptions of our bundle, say 'the man called "Aristotle,"' 'the tutor of Alexander,' 'the master of those who know' and 'the lover of Herphyllis.' This changes absolutely nothing in our judgment! Although his name was also 'Aristotle,' he could not be the greatest philosopher of ancient Greece. He could not be – even though he did in fact educate his son Alexander – because this son could not be the greatest conqueror of Antiquity. Even if someone called him 'the master of those who know,' it does not matter, as that person would surely not be Dante Alighieri. And he could not, even if he had a mistress named Herphyllis, have a relationship with a concubine from ancient Stagira. It does not matter how many auxiliary conditions related to our true Aristotle this proper name satisfies, they will never suffice to identify him. We regard them as unforeseeable irrelevant, strange coincidences, showing that usually auxiliary descriptions can only be helpful – as we will see – insofar as fundamental descriptions are already applicable.[6]

[6] In this connection, we could ask about the role of causality. Though I presently believe that causality must play an inevitable role in naming – we couldn't have real naming without any place for causality – this role is usually *presupposed*. Imagine the explanation of a fortune-teller who, after gazing into a crystal ball, always correctly guessed the name of unknown new visitors… If we believed in something beyond stage magic, we might assume the seer had some sort of occult gifts, e.g., clairvoyance, which explain her ability to name visitors before hearing

3. Objection of vagueness

At this point, one could object that identification rules derived from the MD-rule (or instantiating it) are too vague. Not only do they not establish precisely how much of an inclusive disjunction must be satisfied in order to be sufficient, they also do not establish precisely how much more a possible competitor must satisfy the disjunction in order to disqualify another with certainty. Moreover, to some extent localizing and characterizing rules contaminate one another…

To answer this question, we need to begin by remembering that vagueness does not mean (as shown, e.g., by the *sorites* paradox) the disappearance of boundaries. After all, it is quite easy to imagine a possible world where we could not be certain if our Aristotle ever lived there. This would be the case, for example, in a near possible world where Aristotelian philosophy never developed, but there was a court physician in Stagira named Nicomachus who in 384 BC fathered an anencephalic baby that died soon after birth. The parents even named their new offspring 'Aristotle'... Would the child be our Aristotle? We cannot tell.

Having this in mind, the objection is easy to answer. Our natural language is vague; for our semantic rules to be truly applicable in other empirically possible worlds, they must leave room for vagueness. This is precisely what our identification rules do. Thus, far from being a problem, their vagueness can be very well justified. For the vagueness of our MD-rule is evidence of its correctness, since all it does is to mirror the semantic vagueness already present in our practice of naming. In our example, this is shown by our bicephalous Aristotle, for whom the condition of sufficiency cannot be met because of its unavoidably blurred borders.

Saul Kripke correctly classified proper names as *rigid designators*, defining a rigid designator as a term that designates the same object in every possible world where this object exists or could exist while designating no object in any world where this object does not exist (1971: 145-6).[7] Excluding the confusing 'could exist,' I see this as an intuitively

---

their names. But this clairvoyance, we assume, should be in some way causally explainable! Suppose now that by an incredible coincidence one dreams of a beautiful snow covered volcano called Ozorno, located in Southern Chile. This volcano really exists, which means that the reference is purely *coincidental*. But is a purely coincidental reference more than just a surprising accidental appearance of reference**?** I am now inclined to answer negatively.

[7] Kripke also calls a rigid designator a term that designates the same object in *every* possible world (even in worlds where the object does not exist) (1980: 49). This would cover cases like that of the contra-factual supposition that Hitler was

useful device for selecting the adequate theory of proper names. From this idea we can derive the following rigidity test: If *x* is a rigid designator, its reference could not have existed without being *x*.[8]

Now, since in some possible worlds there are cases where we cannot know whether the bearer of a proper name exists, we must redefine the rigid designator as the term that applies in every possible world where its reference *definitely* (unambiguously) exists. So understood, the proposed MD-rule again makes the proper name a rigid designator – a point to which we will return below.

4. Signification

The proposed meta-descriptivist analysis of how proper names work can help us gain a better understanding of proper names' *meanings*.[9] Leaving aside for now what individual speakers can mean with a name – which is variable and often very limited – we can say that the *core meaning* (sense) of a proper name is given by its fundamental description-rules. The reason for this is that together with the entire identification rule they are able to single out the name's meaning by identifying its sole bearer. These fundamental descriptions must be conventions that are sufficiently known, at least by what we may call *privileged speakers* (in many cases of so-called 'specialists,' these conventions may even be only complementarily shared among them…), understood as those who (alone or jointly) are truly able to apply them in making an identification. So, if you know that Aristotle was 'the philosopher who wrote the *Nicomachean Ethics* and the

---

never born, in which case the name 'Hitler' would refer rigidly in a possible world where that dictator never existed. However, since I cannot make this suggestion less obscure than it is, I evaluate it more as a dispensable work of conceptual contortionism.

[8] I adopt this from Christopher Hughes (2004: 20).

[9] There are several decisive arguments to counter the old objection (Ziff 1960: 93-94) that proper names have no meaning. For one thing, identity sentences like 'Mt. Everest is Chomolungma' are informative, which means the names must have different senses. Moreover, proper names refer to entities belonging to specific classes, e.g., Mt. Everest is a mountain and not a prime number. And the existence of proper names' references can also be negated. Thus, the statement 'Vulcan does not exist' suggests that the name Vulcan must at least have a meaning since the statement does not deny the existence of the word 'Vulcan.' (*Cf.* Searle 1969: 165 f.) My own view is that proper names have so much meaning unevenly distributed in their different cases of application that at first glance they appear to have no identifiable meaning at all.

Metaphysics' and that he was 'a person born in Stagira in 384 BC, son of the court physician Nicomachus, who lived most of his life in Athens,' you already have some decisive informational meaning, though not if you only know that he was 'the tutor of Alexander.'

Now, what about auxiliary descriptions? I think they are still able to give an *aura* of meaning to a name, which sometimes becomes very suggestive, as in the case of 'Plato's greatest disciple.' Nonetheless, someone who only knows a supplementary description usually associated with a proper name, like 'the teacher of Alexander' in association with 'Aristotle,' whom he saw portrayed by an actor in a movie about the life of Alexander, does not really know anything relevant about the meaning of the name Aristotle. Yet, even if he only has a little meaningful information about him, he can use this meager knowledge to insert the name correctly in some sort of vague discourse, producing a *parasitical* kind of reference. Auxiliary descriptions have an auxiliary role of guiding a speaker within a linguistic community, opening an informational channel that directly or indirectly links him on a chain to other speakers, ending, he supposes, with those speakers who definitely know the identification rule, and could in principle teach him to properly identify the bearer (this has sometimes been called a process of 'reference borrowing').

Finally, here we should not confuse cognitive with emotive meaning. The bundle of descriptions associated with a proper name, particularly regarding fundamental descriptions, gives its informative or cognitive content – what Frege called its sense (*Sinn*). This has a conventional ground that is in some way implicitly or explicitly established as something that can be shared among the speakers. However, there are also things like images, memory-images, feelings, smells … that can be strongly associated with a proper name (e.g., 'the Pietà,' 'Gandhi,' 'Stalin,' 'Auschwitz'…), but cannot be easily captured by descriptions. We could say they belong to an imagistic-emotive dimension of meaning, which would be based on the often-shared regularities of our psychological reactions instead of our usually implicitly established conventions. The widely disseminated idea that *not all* our cognitive meanings can be linguistically expressed in the form of descriptions appears to have arisen from a failure to distinguish imagistic-emotive senses from conventional meanings. Because it is descriptively expressible and, regarding ordinary human language, conventionally grounded, cognitive meaning has a shared basis that allows hearers to decode it.

5. Ignorance and error

Possessing this general explanation of the meaning of proper names, we are prepared to give an answer to Kripke's counterexamples of ignorance and error. They concern people who associate an *indefinite* description with a proper name, such as 'a physicist or something like that' with the name 'Feynman.' They also concern people who associate *erroneous* descriptions with a proper name, such as 'the inventor of the atom bomb' with the name 'Einstein' or 'the originator of Peano's axioms' with the name 'Peano' (actually these axioms were first conceived by Dedekind and later refined by Peano) (Kripke 1980: 81-89).

My answer is that the speaker is already able to endow proper names like these with a merely parasitical or borrowed referential role. To do this, it suffices to know a very marginal or relatively inadequate description, as long as one has reasons to believe that in the linguistic community the name has a reference supported by privileged speakers with the necessary knowledge of the fundamental description-rules that enables them to apply the identification rule. This means that a speaker who only knows such insufficient descriptions is already able to insert a word into the discourse in a way he expects can associate the name with its bearer somewhere in the communication network. Important for the success of this parasitical form of reference is that the description known by the speaker enables him to insert the proper name in an understandable way *into sufficiently vague discursive contexts*. This is the case of the Kripkean counterexamples presented above. One can correctly insert names in a sufficiently vague discourse by associating them with even just one indefinite or erroneous description, insofar as at least the following two conditions for parasitical reference are satisfied:

(A) The description known by the speaker must be *convergent*. That is, a description that at least correctly classifies the name's owner (e.g., belongs to class C of the name's identification rule).
(B) The speaker implicitly *knows the MD-rule* for proper names. This means he must be well-aware that he does not know more than an irrelevant part of the meaning, which will make him sufficiently cautious about inserting the name in discourse (he knows how little he knows).

To give a simple example: Not being a theoretical physicist, I know very little about the cognitive meaning of the abstract name 'string theory.' But at least, I am aware of how little I know. This is why I could even impress

my students by giving some vague information about super-strings as incredibly small vibrating filaments of energy that produce all the matter and energy in the universe by vibrations of different frequencies… The ordinary context allows this, although in fact I am far from understanding the relevant mathematical concepts and equations constitutive of the theory. This is why I would refrain from participating where expert knowledge is required, for instance in a discussion among theoretical physicists. Furthermore, without real privileged speakers and their adequate knowledge of meaning, my insertion of the word 'string theory' into discourse would be vacuous, for this parasitical reference borrowing would, in the end, have nothing to anchor itself to. If all specialized knowledge of string theory should disappear due to a cosmic catastrophe that killed all the string theorists and destroyed all their scientific instruments, texts, and data, the real cognitive meaning of this name would also be lost, even if someone could still remember how to pronounce it.

Consider now Kripke's counterexamples. A person can insert the name 'Feynman' in sufficiently vague discourses. His use must be convergent, he must correctly classify Feynman as 'a physicist or something like that' and therefore as a human being, and he must be implicitly aware of the MD-rule. A person can also use the names 'Einstein' and 'Peano' correctly in vague discursive contexts, possibly expecting to obtain more information or even correction from better informed speakers. He must simply satisfy conditions (a) and (b), correctly classify Einstein as a scientist, Peano as a mathematician and both as human beings…

On the other hand, when proper names are associated with *divergent* descriptions, that is, incorrectly classified, the referential thread is apt to be lost. Thus, if speakers associate the name 'Feynman' with the divergent description 'a brand of perfume,' the name 'Einstein' with the divergent description 'a precious stone,' and the name 'Peano' with the divergent description 'a musical instrument,' they will probably not be able to insert these names correctly in any discursive context, no matter how vague it may be. We will not say that in using the name they are able to refer to its bearer, even in an assumed borrowed or parasitical way.

Curiously enough, the same conditions also apply to general terms. If a fisherman means by a whale a large marine fish, this is incorrect, as whales are mammals, but at least it is convergent since he classifies the whale correctly as a sea creature, which already enables him to insert the word in colloquial discourse. However, if a child believes that 'whale' is the name of a mountain in the Appalachians, his usage is not only incorrect but also divergent, making him unable to adequately insert the word into discourse.

Finally, I can use what we have learned to refute a counterexample to descriptivism suggested by Keith Donnellan (1972: 374). He describes a case in which a close friend, Tom, visits a couple and asks to see their child, who is asleep in his bed upstairs. The parents agree to his request, awaken the child and introduce their friend, 'This is our friend Tom.' Tom greets the child with 'Hello,' and the child, after hearing this, immediately falls asleep again. Asked about Tom the next morning, the child replies, 'Tom is a nice person,' without associating any definite description with Tom. Even though he would most likely be unable to recognize Tom on other occasions, according to Donnellan he has still succeeded in referring to Tom!

The answer can vary depending on the details of the story. If the child has no memory of being awakened, of having seen or heard anyone, then he is only trying to satisfy his parents. In this case, of course, he is not actually referring to anyone. However, let us suppose that the child still has some vague memory of seeing a strange person the previous night. If he saw Tom on the street, he would not recognize him. Nevertheless, in this case, he is already using the proper name in a *convergent* way since he associates the name 'Tom' with the description 'a friendly person I saw last night.' In this particular discursive context, hearers who know the identification rule for Tom will be able to give the utterance its full meaning. The parents are privileged speakers here. They know Tom's appearance, what he does for a living, where he lives, where he comes from and many other details of his life. Indeed, without this additional knowledge, the child's comment would be empty, not really being elucidative as a borrowed way to refer to a particular Tom in any satisfactory sense. The child's vague description must be supplemented by his parents, who arranged for Tom to meet the child last night (an adventitious description), know the causal circumstances and are able to refer to Tom in the full sense of the word by means of his name's identification rule.[10]

[10] I am not the first to perceive the inadequacy of this counterexample. As one commentator wrote: 'But why then say that he has in the appropriate sense referred to someone? The language we use to describe such cases may be misleading. Suppose, for example, that Jones is trying to remember whether he has invited someone besides A, B and C to dinner; he has the feeling that there may be one or two more. One might say to him: "Whom might you have in mind?" The point is that in other rather more central uses of these expressions, there is in this sort of case no one whom he has in mind, or means, or refers to; he cannot remember.' (Brian Loar 1976: 367)

6. Rigidity

The proposed meta-descriptivist view explains why proper names are rigid designators, namely because their identification rules apply in any possible world where the proper name's bearer exists. They must pass the rigidity test: the reference of an identification rule *x* could not exist without *x* being applicable to it. It is easy to find a descriptivist explanation for this. What this really means is that a name's bearer, the object, cannot exist without satisfying its identification rule, since this rule simply *defines* what this bearer can be in any world-circumstance. Once established, an identification rule is able to generate all the possible combinations of descriptions of particularized properties (tropes) that an object must have in order to be the sole bearer of its proper name. Consequently, a proper name's identification rule *necessarily* applies to its object of application, if this object exists.

To clarify this point we can express a proper name's identification rule in the form of a *definitional* identity sentence able to single out the bearer of the proper name by means of a complex associated definite description. As an example, we need only formulate the rule identifying what we mean by the proper name 'Aristotle' with what we mean by the following complex definite description:

> Aristotle (*Df.*) = *the* name that in any world-circumstance where it has a bearer applies to a human being who sufficiently and more than any other satisfies the condition that he was born in Stagira in 384 BC… died in Chalcis in 322 BC and/or... was the author of the relevant views belonging to the Aristotelian opus.

Rightly understood, this identification is an analytically *necessary a priori* statement. It contains the complex definite description '*the* name that in any possible world where…,' which besides defining what the name means is a rigid designator. It passes the proposed rigidity test: the name 'Aristotle' is rigid because one cannot imagine a possible world where Aristotle exists but is not Aristotle because he doesn't satisfy the above definition; if Aristotle exists, the definite description necessarily refers to him.

We can see that, unlike the old descriptivism, the meta-descriptivist view does not risk destroying the rigidity of proper names. Quite to the contrary, it enables us to show their rigidity descriptively, since it explains the conditions under which any possible world may be home to the bearer of a name, to whom the name necessarily applies. The reference occurs by

means of particularized properties or tropes (See Appendix to Chapter III), insofar as they satisfy criterial configurations that can be generated by the rule and are seen as sufficient for its application. However, the particularized properties that satisfy the respective criterial configurations do not need to remain the same. They can change in multiple and varied ways, constituting no permanent individualizing essence. Thus, in one possible world we can identify Aristotle as a person born in 384 in Stagira as the son of Nicomachus, and in another possible world we could identify Aristotle as the person who wrote the *Metaphysics* and the *Organon*… This flexibility of the identification rule frees us from having to include essential properties of the referred to object that must be seen as necessary and sufficient conditions for the name's application; the essence belongs here rather to the rule's functional structure. It could be rather called a 'nominal' essence.

7. Rule changeability

One objection is that the changeability of conventional rules would destroy rigidity. Consider the following supposed counterexample. It is well-known that earlier in his life – a period called Δt_1 – David Hume was known as a major historian but not as a philosopher, and thus our characterizing rule for him could be 'the author of *The History of England.*' I call this corresponding early identification rule IR-Hume1. Now, at a later time and up until the present – a period called Δt_2 – Hume became much better known as 'the author of the *Treatise*' rather than as 'the author of *The History of England.*' I call this present identification rule containing more information IR-Hume2. Now, suppose that at a future time Δt_3 the information that Hume was the author of *The History of England* is completely lost, and the only remaining characterizing description is 'the author of the *Treatise*.' I call the identification rule containing this characterizing rule IR-Hume3. Comparing IR-Hume1 with IR-Hume3, we see a case in which the available characterizing identification rule completely changes. Now imagine there is a possible world *Wr* where there is a single Hume who only wrote *The History of England* and another possible world *Ws* where there is a single Hume who only wrote the *Treatise*. In this case, we would apply IR-Hume1 to the historian of the *Wr* and IR-Hume3 to the philosopher of *Ws*, perhaps identifying different persons in the different worlds.

Now, imagine a bizarre situation. Suppose that in a possible world *Wt*, very similar to ours, there were identical twins, the Humes, who had insufficiently different localizing descriptions, but one was only the

historian, while the other only wrote the *Treatise*. Now, using the rule IR-Hume1, we would identify the first person as our Hume. However, using IR-Hume3, we would identify his twin as our Hume. This seems sufficient to show that enough change in the identification rule can lead us to identify different objects in the same possible world, destroying the rigidity of the proper name.

My first reaction to this result is to concede that the right way to preserve rigidity is to agree with the following *condition of preservation*:

> CP: A proper name's owner must be what is meant according to a single identification rule established by privileged speakers of a language community at some Δt.

To this extent, at least, rigidity is warranted. This is the case of our own IR-Hume2, which due to the condition of predominance identifies the writer of the *Treatise* in *Wt* as our Hume and not his twin, since the latter only wrote, less relevantly, *The History of England*. It is also important to add that CP is applicable to fundamental description-rules like those characterizing Hume. Auxiliary descriptions can always change without affecting rigidity, since they are non-definitional.

It is worth noticing that comparing IR-Hume1 with IR-Hume2 we see in the latter an *increase* in the number of details of meaning, making the rule more complex. With IR-Hume2 we have more elements with which to identify the same object, and we would have more resources to identify the same object in possible worlds. We could, I suppose, identify it in possible worlds where we couldn't *definitely* identify the object by using IR-Hume1 alone.

The example of a transition from IR-Hume regarding only the historian, like IR-Hume1, to IR-Hume2, is important because it is usual: normally our information about a proper name's owner increases with time (think of very detailed biographies and autobiographies), what might include fundamental descriptions. That is, over time we usually add new descriptions to a normally unchangeable core, making the boundaries of its application sharp enough to decide on doubtful cases that earlier lay within the blurred borders of application (some possible worlds where the applicability of the name was undecidable are now decidable). However, this is not sufficient to destroy rigidity, since because of it, we do not choose different objects in possible worlds where the object exists, but only improve the acuity of our identification. Hence, I conclude that we are allowed to add a complementary condition of conservation to CP:

> CP1: If we accept change and increase in the details of identifying conventions by the privileged speakers of a language community at some Δt without altering their relevant nucleus of meaning, this does not force us to abandon rigidity.

For now, this suffice. However, it is relevant to note that the amplification and change of descriptions associated with a proper name would cause a real trouble for Kripke's view, albeit hidden by his coarse-grained analysis. For by what means could he identify the right Hume in *Wt*, except by implicitly regarding him first of all, as 'the writer of the *Treatise*,' namely, the person who satisfies our IR-Hume2?[11]

An example that helps to explain the point is that of the island now known as 'Madagascar,' suggested by Gareth Evans as a possible argument against Kripke's causal-historical view of proper names (Evans 1973). 'Madagascar' was initially used as the name for the eastern regions of Africa. During his world travels, Marco Polo visited a large island off the coast of eastern Africa and mistakenly began to use the name Madagascar for it. Today, because of Marco Polo's mistake, we all use to call this island Madagascar. However, if Kripke's theory were correct, according to which the reference of a name is fixed by a causal-historical chain beginning with its first baptism, we should still use the name Madagascar for the eastern part of Africa. Kripke tried to solve this problem by suggesting that there is a new social intention to refer to the island that *overrides* the former intention (Kripke 1980: 163). However, this answer dangerously approaches a recognition of the necessity of new descriptions (disguised as intentions) to identify the island.

From our perspective, we can easily solve the problem. We could admit that this is a case of homonymy, since there is a forgotten Madagascar-1 of eastern Africa, with its proper fundamental description-rules, and the well-known Madagascar-2, the island, with very different proper fundamental description-rules. Here we have a new identification rule created for a new reference using the same proper name's symbolic form. We have a complete change in the nucleus of meaning, which precludes the application of P1.

8. Names *versus* descriptions

Perhaps the decisive advantage of my proposal is that it gives the only really satisfactory explanation of the contrast between the *rigidity* of proper names and the *accidentalness* (*flaccidity*) of definite descriptions.

[11] For the examination of a real case, see Costa 2014: 51.

According to Kripke, unlike proper names, definite descriptions can have different bearers in different possible worlds. So, while the name Benjamin Franklin always refers to the same person in any possible world where this person exists, the description 'the inventor of bifocals,' which refers to him in our world, could refer to a different person or even to no person in some other possible world. In Kripke's case, this is because proper names, being rigid, are after their bearers' baptism necessarily linked with them in a mysterious way that calls for explanation. Definite descriptions belong to a different epistemic category. We could say that they refer by means of what J. S. Mill would call their *connotation*, defined by him as the implications of attributes belonging to the object referred to (Mill 2002: 19, 21).

Nonetheless, in making this sharp distinction Kripke overlooked the most relevant point, namely, that *definite descriptions are only accidental when associated with proper names.* The point can be made clear first intuitively and then using Wittgenstein's distinction between symptoms and criteria.

Intuitively, the reason why most definite descriptions are accidental designators (such as 'the inventor of bifocals') *is that when we apply them we integrate them semantically, in a contingent way, with the identification rule of some proper name* (such as 'Benjamin Franklin'). Indeed, this integration isn't normally established as necessary by identification rules (and our MD-rule). Consequently, we can easily imagine possible worlds where there is a *mismatch* between the object possibly referred to by a proper name and the object possibly referred to by the definite descriptions usually attached to it, particularly when these descriptions are merely auxiliary ones (for instance, in a world where Samuel Adams invented bifocals and Benjamin Franklin never existed).

The foregoing explanation of the distinction between the rigidity of proper names and the accidental character of descriptions can be elaborated with the help of Wittgenstein's distinction between *symptoms* and *criteria*. According to this distinction, once accepted as given, a criterion *warrants the application of a word*, while a symptom, once accepted as given, *makes this application only more or less probable* (See Ch. II, sec. 9 and Ch. III, sec. 10 of this book). In their association with proper names, definite descriptions usually give us only symptoms for their application, particularly when they are auxiliary, though sometimes even when they are fundamental. This explains why these descriptions alone are not applicable in all possible worlds where the bearer of a proper name exists. By contrast, the complex definite description expressing the whole identification rule of a proper name is able to generate multiple

independent *criteria* to identify the referent (e.g., Aristotle) in different possible worlds. These criteria can be met by particularized configurations of properties (understood as *tropes*) like those satisfying the examples given above. One can say that in different possible worlds the bearer of a proper name can satisfy the same identification rule in different ways, by means of many different possible configurations of particularized objective properties/tropes.

An easy way to prove that my reasoning is correct is by explaining a phenomenon that Kripke's causal-historical view cannot explain. We only have to find definite descriptions that are *not* semantically associated with any proper name. In this case, we expect them to behave as rigid designators, applying to only one object in any possible world where this object exists. I call them *autonomous* definite descriptions. The following four descriptions are examples:

1. the 52nd Regiment of Fot,
2. the last living Neanderthal,
3. the 1914 assassination of Austrian Archduke Ferdinand in Sarajevo,
4. the easternmost point of South America.

These descriptions respectively name a military organization, a human being, an event, a place. What matters is that they are all easily recognized as rigid designators. Consider:

(1) We can imagine a similar possible world where the 52nd Regiment of Fot had a different organization and time of existence, for example, a world where it did not serve in the Napoleonic wars.
(2) We can imagine a possible world where the last Neanderthal outlived all members of the species *homo-sapiens*.
(3) We can imagine a world where the Archduke was assassinated at a different time by someone other than Gavrilo Princip and by other means.
(4) And we can imagine a possible world where the easternmost point of South America is not in Brazil but in Tierra del Fuego, which in this world stretches far eastward towards Africa (assuming that we are considering 'the same point' regardless of determined properties and latitude).

Even so, if applicable these descriptions will always be applicable to *the same bearer* in every possible world where this bearer exists, whether it is

an organization, a human being, an event or a geographic location. These definite descriptions are rigid designators simply because, with their localizing and/or characterizing description-rules, made at least partially explicit by them, they are always able to pick out the same referent, without the danger of mismatching with referents picked out by the identification rules of associated proper names. The Kripkean view would have no explanation for this, except by an *ad hoc* claim that autonomous descriptions are nothing but disguised proper names.

9. Autonomous definite descriptions

Finally, it is worth noting that the same MD-rule we apply to bundles of descriptions of proper names can be applied in the case of autonomous definite descriptions, insofar as they work as rigid designators and singular terms independent of any proper name. The difference is not just that part of the rule is usually made explicit through symbolic forms (as a 'connotation'), but that the rule is often less complex. I can give as an example the identification rule for 'the 52nd Regiment of Fot.' It has the following (summarized) localizing description-rule:

> The 52nd Regiment of Fot existed from 1757 to 1881, stationed in Oxford; it saw active service particularly during the American War of Independence, the Anglo-Mysore wars in India and the Napoleonic Wars.

The identification rule for the 52nd Regiment of Fot has the following (summarized) characterizing description-rule:

> The 52nd Regiment of Fot was a highly regarded regiment whose troops were recruited chiefly from Oxfordshire, consisting of one or two battalions of light infantry, each comprising approximately 1,000 men.

Of course, the inclusive disjunction of these descriptions needs to be only sufficiently and predominantly satisfied in any possible world-circumstance where 'the 52nd Regiment of Fot' exists. Auxiliary descriptions are also present, for instance 'the regiment never surpassed in arms, since arms were first borne by men,' though they are of lesser relevance. The same is the case with other autonomous definite descriptions.

On the other hand, most definite descriptions, like 'the inventor of bifocals' or 'the tutor of Alexander' or 'the City of Light' (*la Ville Lumière*), are employed in close association with proper names (respectively Benjamin Franklin, Aristotle, Paris). In this case, the descriptions are viewed as merely auxiliary ones, emphasizing their explicit connotations. As such, they are seen as complements to the identification rule of their associated proper names.

10. Kripke's counterexamples

The above exposed meta-descriptivist theory of proper names demonstrates its explanatory power when we need to refute standard counterexamples to descriptivism. As we noted, the meta-descriptivist rule is a tool that all competent users of proper names must be able to use, even without being aware of it. Having made this tool explicit, we can now rehabilitate descriptivism by giving more satisfactory answers to objections and counterexamples. We already saw this in its capacity to answer Kripke's modal objections, according to which descriptivism is condemned because any description or group of descriptions associated with a name can fail to apply to the name's bearer, while as rigid designators proper names never fail to refer to their bearers. To justify this view further, I will first consider Kripke's main counterexamples.

(i) The first is Kripke's memorable Gödel counterexample (1980: 83-84). Suppose Mary knows nothing about Kurt Gödel, except the description 'the discoverer of the incompleteness theorem.' Then suppose that in nineteen-thirties Vienna an unknown Viennese logician named Schmidt wrote the first paper to describe the incompleteness theorem but died under mysterious circumstances before he could publish this major discovery. Soon after this his friend Gödel stole his manuscript and published it under his own name. According to Kripke, if the descriptivist theory were correct, Mary should conclude that Gödel is Schmidt. But it is obvious that the name 'Gödel' still refers to Gödel and not to Schmidt! And according to Kripke, the reason is that the reference is fixed by the baptism of the infant Gödel. This is followed by a causal-historical chain in which each hearer repeats the name with the intention to refer to the same person referred to by the speaker from whom he heard it, continuing up to Mary's utterance…[12]

[12] Note that this intention should have no proper cognitive content, otherwise we would be able to express this content linguistically, falling back into descriptivism.

However, this objection poses a threat only to Kripke's own caricatured formulation of descriptivism. Our identification rule for the name 'Gödel' goes much farther. First, the characterizing description-rule for the name 'Kurt Gödel' can be summarized as:

> a great logician who made major contributions to logic, particularly the incompleteness theorem.

This already indicates more than what Mary knows, since this description also points to Gödel's other contributions to logic. Moreover, Kripke does not even consider the localizing description-rule, which can be summarized as:

> the person born in Brünn in 1906 who studied in Vienna, emigrated to the USA in 1940 via the trans-Siberian railway and worked at Princeton University until his death in 1978.

As a competent speaker of the English language, Mary must implicitly know the MD-rule. She must be tacitly aware that to conclude that Gödel was Schmidt, she would have to do much more than just attribute the discovery of the incompleteness theorem to Schmidt. Consequently, she wisely refrains from concluding that Gödel is Schmidt.

Moreover, for a privileged speaker Gödel cannot be Schmidt, because even if Schmidt satisfies part of Gödel's characterizing description, Gödel continues to satisfy the whole localizing description and at least part of the characterizing description, satisfying in this way the condition of predominance. Nevertheless, we already can see that something in the meaning of the name 'Gödel' is attached to the name 'Schmidt,' which would be clear if someone heard a mathematician who, scandalized by this information, resorting to hyperbole, angrily protested: 'No! The true Gödel was Schmidt!'

Furthermore, under certain circumstances, Gödel could really be Schmidt. Suppose that Schmidt killed Gödel when he was a teenager and assumed his identity. Then Schmidt studied mathematics in Vienna, conceived and published the incompleteness theorem, married a woman named Adele, moved to the USA in 1940 via trans-Siberian railway and worked at Princeton University until his death in 1978. In this case, we would all agree that Gödel was, in fact, Schmidt, the unscrupulous

---

A problem is that without proper content this intention would be nothing but a *desire*, a *bet* on the sameness of reference.

murderer. And the famous photo of Gödel with Einstein would actually be a photo of Schmidt with Einstein. But why should we say this? The answer is clear: because we see that Schmidt now satisfies the condition of *predominance*. Smith now sufficiently satisfies our localizing and characterizing description-rules for the name 'Gödel' much more than the unfortunate teenager whose birth-name he stole. And since his true birth-name was 'Schmidt', he also satisfies the localizing and characterizing identification conditions of Schmidt before he murdered Gödel. The identification rule we now attach to the name 'Schmidt' includes the great majority of conditions constitutive of the rule we earlier attached to the name 'Gödel.'

(ii) Now consider the case of semi-fictional names like Robin Hood. From my perspective, if a name really is semi-fictional, it must be associated with some descriptive content effectively applicable to a real owner, along with merely imaginary descriptive content added later, even if we are unable to definitely distinguish the first type of content from the second. – If they lack any descriptive content that we could consider applicable to reality, they should be called 'purely fictional names.' Thus, with regard to semi-fictional names, in many cases, our situation is one of uncertainty and insufficient knowledge. This is the case of Robin Hood. The vague descriptions 'a person who probably lived in England in the 13th century' and 'a legendary righter of wrongs' respectively suggest contents of almost completely unknown localizing and characterizing descriptions.

According to Kripke, the story is different. It does not matter whether a semi-fictional name has any true descriptive content. Important is only that the name meets its own requirement of coming at the end of the right external causal-historical chain linking it with the baptism of its reference. Hence, independently of any bundle of descriptions known or unknown to us, if this condition is met, the reference of a semi-fictional name is warranted.

Our descriptivist answer is more balanced and complete. As descriptivists, we should admit that what we think is a semi-fictional name can, in fact, be purely fictional.[13] We suspect that the name has a reference, since there are hints that it could refer to a real historical person, so that one could find its proper fundamental descriptions at least in principle. For instance, suppose historians discover documents about a

[13] The supposedly semi-fictional name originally used as an example by Kripke was that of the biblical prophet Jonah (1980: 67-68). However, the majority of serious Bible scholars believe that Jonah was, in fact, a purely fictional character.

man named Robart Hude, an early 13th century outlaw who championed the weak against the powerful and lived in hiding with a band of followers in Sherwood Forest near Nottingham, strongly suggesting that his life story may have given rise to the legend of Robin Hood. With this in mind, we have enough information to apply both the correct localizing description – early 13th century, lived near Nottingham – and the correct characterizing description – an outlaw who stole from the rich and gave to the poor – both summarized fundamental descriptions originating the legend of Robin Hood. This would give us an improved descriptivist confirmation of the origin of Robin Hood as a confirmed semi-fictional character, while a causal-historical 'explanation' should not change anything.

We can also find cases suggesting inadequacies of Kripkean explanations. A scholar might discover there really was a historical model for the first medieval author who wrote about the legend of Robin Hood, but that none of the traditional descriptions apply to it. Suppose there was a faithful hunting dog called Robin who tagged along when the medieval author went hunting in Sherwood Forest. Inspired by the loyalty and bravery of the dog, who was always ready to help his master, the author created the fictional story now known as that of Robin Hood. In this case, it seems that a Kripkean philosopher should conclude that Robin Hood was the dog's name. A historical chain began when the writer baptized the puppy, and all subsequent hearers or readers shared an intention to refer to the same subject as Robin Hood, though adding the most diverse descriptive fantasies. But this certainly strikes most readers as more than a bit strange.

On the other hand, our MD-Rule allows us to explain the case more clearly and persuasively. This rule would indicate that Robin Hood was the name of a purely fictional character and has nothing to do with any dog since according to the identification rule, the bearer of the name 'Robin Hood' should at least belong to class C of human beings.

(iii) The most elusive counterexample is that of Hesperus (1980: 57-58). Suppose, says Kripke, that someone once fixed the reference of Hesperus by using the following statement (i) 'I shall use "Hesperus" as *the* name for the heavenly body appearing in that particular position in the sky.' This does not mean that to be in a certain position in the sky is a necessary property of Hesperus. If long ago a comet had collided with the planet Hesperus, it might no longer have been in its original position when first discovered. Nevertheless, Hesperus would still be Hesperus, since the

name is a rigid designator. For Kripke, the bundle theory is unable to explain this case.

Our answer comes from an analysis of the identification rule for Hesperus (the Evening Star). Although one can naively define it as the most brilliant celestial body in the evening sky, one can also call 'Hesperus' the planet Venus because it has always appeared to us as the Evening Star. Moreover, it is clear that with the word 'Hesperus' Kripke had the planet Venus in mind. If Hesperus were meant to refer not to Venus, but only to 'the luminous thing over there,' Kripke could not speak of a comet colliding with the planet, etc.[14] This considered, the identification rule that Kripke in fact applies is that of Venus, even if by chance appears to us as the Evening Star. (An astronomer can point to the evening Star and say: 'over there is Hesperus, the planet that is so bright because it is covered by highly reflective clouds of sulfuric acid,' meaning the planet Venus in general and not so much its accidental appearance as the Evening Star). The identification rule for the planet Venus (identified by the appearance of Hesperus) as really considered by Kripke can be summarized as follows:

> IR-Venus: In any world-circumstance where there is a bearer of what we call with the proper name 'Venus', this bearer must be: (i) a celestial body that (ii) sufficiently and (iii) more than any other satisfies (iv) the condition of being the second planet from the Sun in the solar system, orbiting between Mercury and Earth as a proper constituent of the system at least for some time. (Moreover, there is a bundle of auxiliary descriptions like '*the* brightest celestial body in the evening sky visible in the direction of the Sun,' which are contingent.)

It must be noted that this is a one-foot identification rule, since the localizing feature of Venus is its being the second planet of the solar system orbiting between Mercury and Earth, and its characterizing feature of being a planet is already included in the localizing description. Moreover, other properties of Venus, insofar as they do not prevent the application of the identification rule, are irrelevant. They are the objects of auxiliary descriptions, one of which is that it appears to us as Hesperus. A planet with all the characterizing properties of Venus which didn't belong to our solar system but rather to a solar system in another galaxy would

[14] After all, the seed of this example was already planted much earlier by R. B. Marcus' example using the name 'Venus' (*Cf.* Marcus 1993: 11; see also Ch. IV, sec. 26 of this book).

surely not be Venus. If Venus, even if called 'Hesperus,' loses its atmosphere and therefore its brightness and cannot be seen anymore, it will remain an inconspicuous Venus. And if Venus loses most of its mass but remains a small planet, it will still be a Venus ('Venus after the storm'). Moreover, IR-Venus is a rigid designator: there can be no possible world where Venus is not the second planet of the solar system… An object can be correctly identified as Venus if it satisfies the localizing description-rule sufficiently and more than any other planet.

Now, suppose that a comet collided with the planet Venus sometime *after* it was identified as the Evening Star, and that this collision changed its position in the sky so that Venus was transformed into an errant planet outside the solar system… In this case, IR-Venus (called 'Hesperus') would remain applicable, even if the description of Hesperus would not be satisfied anymore.

However, Kripke's example is more sophisticated. He invites us to imagine a possible world *Wh* (conceived as our own world under counter-factual circumstances) with no evening star and no second planet between Mercury and Earth in the relevant historical time period. However, – we must suppose in order to make any sense of what he says – astronomers on *Wh* (not necessarily on the Earth…) have discovered that there once was a second planet orbiting between Mercury and Earth belonging to the system, even if it was not visible from the Earth as the Evening Star… but this planet was struck by a comet and was turned into an errant planet or no longer exists… However, this planet still satisfies IR-Venus (…the second planet of the Solar System, located between Mercury and Earth), it can still be called 'Venus' and a would-be 'Hesperus' in Kripke's deceitful use. We see that Kripke's Hesperus, more literally the planet Venus, can satisfy its identification rule even in the possible world *Wh*, since it has as its identification rule a rigid designator, namely a rule that defines all that we can literally call 'Venus' and in some analogical, almost abusive way, also call 'Hesperus,' understood as 'the Venus that by chance appears to us as the Evening Star.'

(iv) Another of Kripke's objections – *circularity* in names like Peano and Einstein – is easy to answer. Limiting myself here to the first, Kripke's view is that we define the name 'Einstein' descriptively as 'the originator of relativity theory,' but we explain relativity theory as a theory authored by Einstein, which leads to circularity.

The answer is not just that it isn't necessarily so (we can explain the theory without mentioning its originator's name), but that the use of its originator's name in its *definiens* is perfectly adequate; for it is natural to

re-utilize a defined *definiendum* in the search for a complete *definiens*. This reutilization is not circular; it is part of an 'ascending bascule movement' in which already available information is used to obtain more information (any Google search should convince you of this).

11. Donnellan's Counterexamples

Now I want to briefly analyze the counterexamples proposed by Keith Donnellan (1970, sec. x):

(i) One instructive counterexample is the following: Suppose, he writes, someone discovers that Thales was actually no philosopher, but instead a wise well digger residing in Miletus, who, despairing of his exhausting work, once exclaimed 'I wish all were water, so I wouldn't have to dig these damned wells.' Now, suppose this sentence came down to Herodotus, Aristotle and others in an altered form as the view of the first Greek philosopher Thales that water is the principle of all things. Donnellan adds to this story the assumption that there really was a hermit who thought all was water. However, he lived in a period so remote that neither he nor his doctrines have any historical significance for us today. We would refuse to accept that the hermit was Thales, even if the hermit really satisfied the description. The reason, according to Donnellan, is clear: Thales and not the hermit was at the start of the causal-historical chain.

The answer offered by our neodescriptivist view is that in some cases the description of a causal history is so important that it must be contemplated in the identification rule. More precisely, it must be included in the characterizing description-rule. This is precisely the case with Thales because what we find *important* about him is that he came at the start of Western philosophy. Without knowing this historical context, the statement 'Water is the principle of all things' would seem ridiculous. Thus, we could summarize the real characterizing definite description belonging to the identification rule for Thales as:

> the person who originated the doxography found in Aristotle, and others, which describes him as having been the first Greek philosopher who said that water is the principle of all things, that everything is alive, etc.

As for the localizing description, we at least know that Thales was:

> the Milesian who lived from 624 to 547-8 BC and probably once visited Egypt.

In view of this, if we return to Donnellan's example we must conclude that according to our version of descriptivism the hermit could not have been Thales! The reason is that Thales the well-digger better fulfills both fundamental conditions, in this way satisfying the condition of predominance.

Let us compare the two cases. The hermit does not satisfy any part of the localizing description; all he satisfies is an incomplete part of the characterizing description. On the other hand, Thales the well-digger completely satisfies the localizing description, because he lived in Miletus from 624 to 547-8 BC. And regarding the characterizing description, even if Thales were not a philosopher and never said the principle of all things is water, he remains the person wrongly described in the doxography as the first Greek philosopher who said all is water. Hence, despite everything, our Thales satisfies the fundamental descriptions much better than does the hermit, thereby qualifying as the name's proper bearer.

Aside from that, one should not forget that depending on whatever details we could add to or subtract from this example, our intuitions would change, leading us to think our Thales never really existed or even that the hermit was the true Thales.

(ii) Another of Donnellan's counterexamples is a student who talked with a person at a party who he believed was the famous philosopher J. L. Aston-Martin, author of 'Other Bodies.' Although the person's name really was Aston-Martin, he only pretended to be the philosopher. Donnellan notes that the sentence (a) 'Last night I spoke with Aston-Martin' is *false* because it associates the name 'Aston-Martin' with the description:

> D1: the philosopher who wrote 'Other Bodies.'

In contrast, the following sentences are *true*: (b) 'At the end of the party Robinson stumbled at the feet of Aston-Martin and fell on the ground' and (c) 'I was almost the last person to leave; only Aston-Martin and Robinson were still there.' This is because they are associated with description D2: 'the man named Aston-Martin whom I met at the party.' The objection is that descriptivist theory does not explain this change. In (a), (b) and (c), the name 'Aston-Martin' should be associated with the same bundle of descriptions that includes 'the author of "Other Bodies."'

The obvious problem with this example is that one can always attach a false description to a proper name, confusing it with a description associated with another person of the same name. The student had the wrong characterizing description, 'the author of "Other Bodies"', and two correct adventitious auxiliary descriptions. He really did not know the identification rule for Aston-Martin. But since he also had (A) convergent descriptions like 'the man called "Aston-Martin"' and (B) an implicit knowledge of the MD-rule for proper names, he was already able to insert the proper name into discourse, even if only to find that he was mistaken.

(iii) A third counterexample suggested by Donnellan is person A, who wearing a special pair of glasses identifies two identical squares on a screen, which are placed one on top of the other. She calls the top square Alpha and the bottom square Beta. The only description suitable for identifying Alpha is its position. Now it turns out that without person A's knowledge the glasses visually invert the square's positions. Actually, Alpha is the bottom square. Donnellan believes he has thus demonstrated that the square which A refers to as the square Alpha is, in fact, the bottom square, even if associating it with the mistaken description:

(a) Alpha = the square that A sees as on top.

In response, I propose that A is insufficiently referring to the square Alpha. She associates the name 'Alpha' with a correct characterizing description (a square) and an incorrect, but *convergent* localizing description, since he still correctly identifies a square as presently given in front of her, hence in the right broad spatiotemporal region. This description is correctable to:

(b) Alpha = *the* square that A sees as on top… even though it is, in fact, the square on the bottom, because A is wearing glasses that invert the positions of the images.

Although observer A does not know description (b), this description is the complete localizing description rule of the Alpha square from A's perspective, as it is known by privileged speakers such as B. Speaker B knows that square Alpha is on the bottom because she has the information expressed by localizing description (b), which gives the referent's mode of presentation. A has a convergent but incomplete and erroneously interpreted description-rule. This is proven by the fact that once she is informed by B about the glasses' inversion of images, she will

immediately replace description (a) with Alpha's true identification rule (b).

12. Explanatory failure of the causal-historical view

Finally, let me say something about the causal-historical view. I do not wish to deny that there is some kind of direct, indirect, or even extremely indirect causal relation between the utterance of the name and its bearer, or some kind of causal-historical relation between the utterance and the first tags of a name's bearer. Even descriptivists like P. F. Strawson haven't denied this. After all, we live in a world of causes and effects, and a proper referential link should have some causal dimension. What I reject is the explanatory relevance of the causal-historical view. No one uses it as a form of explanation. If someone asks me who Aristotle was, I do not answer: 'All you need to do is to continue following my causal-historical chain, without forgetting to keep in mind your intention to refer to the same Aristotle I refer to.'

Indeed, in themselves the causal-historical links will remain inscrutable unless in searching for them we appeal to something like correlative cognitions and consequently to descriptions representing these cognitions. Suppose we had, for instance, an advanced brain scanner able to show that whenever a speaker says the name 'Aristotle' and really knows whom he is speaking about, a recognizable neurophysiological pattern arises in his brain. We could identify this pattern as a link of the external causal-historical chain and search for similar links in other speakers. But since in this case we would need to appeal to the speaker's cognitive-intention, implicitly we would be appealing to descriptions. This shows that the advocate of the causal-historical chain as the only proper explanatory principle commits a *petitio principii* by presupposing descriptivism. To make things worse, Kripke's view of baptism is magical, since it cannot really be based on any property of the referent – it testifies to a form of referential mysticism that blocks the ways of inquiry.[15] Indeed, if we pick out some property, we will have a thought or intention, and consequently this can be in principle descriptively translated. Although philosophically

[15] Defending non-descriptive senses as *mental files* (small packets of information), François Recanati accepts the suggestion that in perception an 'object' without properties could arguably be imagined (2012: 29-31). This might seem true if you expect to identify things like material objects or natural kinds. But when you think about very basic properties like density, hardness, volume, form, color, warmth… something must be present (or even absent in a context of things that are present). Without properties like these, no object can even be hinted at.

original and challenging, as is much of Kripke' work, if taken at face value the causal-historical view of proper names can be reduced to a philosophical fantasy that begs the question. As H. L. Mencken noted, for every complex problem there is always a clear and simple answer that is unequivocally mistaken.

CHAPTER II

THE MOST SUITABLE METHODOLOGY FOR CONCEPTUAL ANALYSIS

> *Eine Art zu philosophieren steht nicht neben anderen wie eine Art zu Tanzen neben anderen Tanzarten ... Die Tanzarten schließen sich nicht gegenseitig aus oder ein ... Aber man kann nicht ernsthaft auf eine Art philosophieren, ohne die anderen verworfen oder aber einbezogen zu haben. In der Philosophie geht es demgegenüber wie in jeder Wissenschaft um Wahrheit.*
> [A way of philosophizing is not one way among others, like one way of dancing among others … Ways of dancing are not mutually exclusive or inclusive … But no one can seriously philosophize in one way without having dismissed or incorporated others. In philosophy as in every science, the concern is with truth.]
> *—Ernst Tugendhat*

> Philosophy has no other roots but the principles of Common Sense; it grows out of them, and draws its nourishment from them. Severed from this root, its honours wither, its sap is dried up, it dies and rots.
> *—Thomas Reid*

Given the commonsense assumptions involved when we take the social role of language as a starting point, at least part of this book must be critical. The reason is clear. The new orthodoxy that dominates much of contemporary philosophy of language is largely based on what I wish to call a *metaphysics of reference and meaning*. Its views often focus on reference more than on meaning, or on something like reference-as-meaning, displaying a strong version of semantic externalism, hypostasized causalism, and anti-cognitivism. I call these views metaphysical not only because they oppose modest common sense but mainly because, as will be shown, they arise from sophisticated attempts to unduly 'transcend' the limits of what can be meaningfully said (*Cf.* Appendixes to Chs. I and II).

One example of the metaphysics of reference is the position of philosophers like Saul Kripke, Keith Donnellan, and others on how to

explain the referential function of proper names and natural species terms. According to them, it is not our cognitive access to the world but rather the mere appeal to external causal chains beginning with acts of baptism that really matters. On the other hand, what we may have in mind when using a proper name is for them secondary and contingent.

Another example is the strong externalist view of Hilary Putnam, John McDowell, Tyler Burge and others, according to whom the conceptual meaning of an expression, its understanding, thought, and even our own minds (!) in some way belong to the external (physical, social) world. Using a metaphor always hinted at but never spelled out, it is as if these things were floating outside, determined by the entities referred to with words, in a way that recalls Plotinus' emanations, this time not from the 'One', but in some naturalistic fashion, from the 'Many.' In writing this, I am only trying to supply the right images for what is explanatorily wanting… In fact, externalism is an unclean concept. After refinements, externalism is defined in a vague way as the general idea that 'certain types of mental contents must be determined by the external world' (Lau & Deutsch 2014). This would be an obvious truism, insofar as we understand the expression 'determined by the external world' as saying that any mental content referring to the external world is in one way or another causally associated with things belonging to an external world. As Leszek Kolakowski once noted, 'if there is nothing outside myself, I am nothing' (2001). But this is trivial enough to be accepted by a reasonable internalist like myself (or by a very weak externalist, which in my view amounts to the same thing). Nonetheless, externalists have proposed in their most central and radical writings to read 'determined' as suggesting that the *locus* of our meanings, beliefs, thoughts and even minds is not in our heads, but somewhere in the external world… However, this sounds very much like a genetic fallacy.

A third example is the view accepted by David Kaplan, John Perry, Nathan Salmon and others, according to whom many of our statements have as their proper semantic contents *structured propositions*, whose constituents (things, properties, relations) belong to the external world alone, as if the external world had any proper meaning beyond the meaning we give to it. As a last example – which I examine in the present chapter – we can take the views of John McDowell and Gareth Evans. According to them, we cannot sum up most of the semantics of our language in tacit conventional rules that can be made reflexively explicit, as has been traditionally assumed. Consistent with causal externalism, their semantic carriers tend to take the form of things that can be understood chiefly in the third person, like the neuronal machinery

responsible for linguistic dispositions unable to become objects of reflexive consciousness.

Notwithstanding the fact that most such ideas are contrary to the semantic intuition of any reasonable human being who hasn't yet been philosophically indoctrinated, they have become the mainstream understanding of specialists. Today many theorists still view them as 'solid' results of philosophical inquiry, rather than crystallized products of ambitious formalist inspired reductionism averse to cognitivism. It is true that they have in the meantime rhetorically softened their extreme views, though still holding them in vaguer, more elusive terms. However, if taken too seriously, such ideas can both stir the imagination of unprepared readers and, more seriously, limit their scope of inquiry.

In the course of this book, I intend to make plausible the idea that the metaphysics of reference is far from having found the ultimate truth of the matter. This is not the same, I must note, as to reject the originality and philosophical interest of its main arguments. If I did reject them on this ground, there would be no point in discussing them here. Such philosophical arguments usually cover insights related to their equivocal conclusions and remain of interest even if they are in the end-effect flawed. If so, they would ultimately require not additional support, but careful critical analysis. In the process of disproving them, we could face views with greater explanatory power, since philosophical progress is very often dialectical. For this reason, we should judge the best arguments of the metaphysics of reference in the same critical way we value McTaggart's argument against the reality of time or Berkeley's remarkable arguments against materialism. Consider Hume's impressive skeptical arguments to show there is nothing in the world except flocks of ideas – an absurd conclusion that was first countered by Thomas Reid. What all these arguments surely did, even if we are unable to agree with them, was to draw illusory consequences from insufficiently known conceptual structures, presenting in this way real challenges to philosophical investigation, useful insofar as they force us to reconsider our views, answering them by means of a more careful analysis of assumed structures as they really are. Indeed, without the imaginative and bold revisionism of the metaphysicians of reference, without the challenges and problems they presented, it is improbable that corresponding competing views would ever acquire enough intellectual fuel to get off the ground.

1. Common sense and meaning

To contend with the metaphysics of reference, some artillery pieces are essential. They are methodological in character. The first concerns the decision to take seriously the so often neglected fundamental principles of common sense and natural language philosophy, respectively assumed by analytic philosophers like G. E. Moore and the later Wittgenstein. According to philosophers with this outlook, we should seek the starting point of our philosophical arguments as much as possible in pre-philosophical commonsense intuitions often reflected in our natural language. The link between common sense and natural language is easy to understand. We should expect that commonsense intuitions – often due to millennia of cultural sedimentation – will come to be strongly mirrored in our linguistic forms and practices.

As Noah Lemos wrote, we can characterize commonsense knowledge as:

> ...a set of truths that we know fairly well, that have been held at all times and by almost everyone, that do not seem to be outweighed by philosophical theories asserting their falsity, and that can be taken as data for assessing philosophical theories (2004: 5).

Indeed, commonsense truths seem to have always reconfirmed themselves, often approaching species wisdom. Examples of common sense statements are: 'Black isn't white,' 'Fire burns,' 'Material things exist,' 'The past existed,' 'I am a human being,' 'I have feelings,' 'Other people exist,' 'The Earth has existed for many years,' 'I have never been very far from the Earth,'… (See Moore 1959: 32-45). Philosophers have treasured some of these commonsense statements as particularly worthy of careful analytical scrutiny. These include: 'A thing is itself' (principle of identity), 'The same thought cannot be both true and false' (principle of non-contradiction), 'I exist as a thinking being' (version of the *cogito*), 'The external world is real' (expressing a realist position on the external world's existence), and even 'A thought is true if it agrees with reality' (correspondence theory of truth).

The most flagrant objection to the validity of commonsense principles is that they are not absolutely certain. Clearly, a statement like 'Fire burns' isn't beyond any possibility of falsification. Moreover, science has truly falsified many commonsense beliefs. Einstein's relativity theory decisively refuted the commonsense belief that the length of a physical object remains the same independently of its velocity. But there was a time when people regarded this belief as a self-evident truth!

This latter kind of objection is particularly important in our context because metaphysicians of reference have made this point to justify philosophy of language theories that contradict common sense. Just as in modern physics new theories often conflict with common sense, they feel emboldened to advance a new philosophy whose conclusions depart radically from common sense and natural language. As Hilary Putnam wrote to justify the strangeness of his externalist theory of meaning:

> Indeed, the upshot of our discussion will be that meanings don't exist in quite the way we tend to think they do. But electrons don't exist in quite the way Bohr thought they did, either. (1978: 216)

One answer to this kind of comparison emphasizes the striking differences between philosophy of meaning and physics: the way we arrive at meanings is much more direct than the way we discover the nature of subatomic particles. We make meanings; we don't make electrons. We find subatomic particles by empirical research; we don't find meanings: we establish them. We do not need to read Plato's *Cratylus* to realize that the meanings of our words are dependent on our shared semantic customs and conventions.

2. Ambitious *versus* Modest Common Sense

I do not have the ambition to end the debates over the ultimate value of common sense. However, I can reasonably demonstrate that two deeply ingrained objections to the validity of commonsense principles are seriously flawed, one based on the progress of science and the other based on changes in our worldviews (*Weltanschauungen*). The first is that science defeats common sense. This can be illustrated by the claim attributed to Albert Einstein that common sense is a collection of prejudices acquired by the age of eighteen… (Most physicists are philosophically naïve.) Changes in worldviews are transformations in our whole system of beliefs, affecting deeply settled ideas like moral values and religious beliefs. In my view, these two charges against common sense are faulty because they arise from confusion between misleading *ambitious* formulations of commonsense truths and their authentic formulations, which I call *modest* ones.

I wish to begin with a closer examination of objections based on the progress of science. With regard to empirical science, consider the sentences:

(a) The Earth is a flat disk with land in the center surrounded by water.

(b) The sun is a bright sphere that revolves around the Earth every 24 hours.
(c) Heavy bodies fall more rapidly than light ones, disregarding air resistance.
(d) Time flows uniformly, even for a body moving near the speed of light.
(e) Light consists of extremely small particles.

According to the objection, it is widely known that science has disproved all these once commonsense statements. Already in Antiquity, Eratosthenes of Alexandria was able to disprove the Homeric view that (a) the Earth is a flat disk rimmed by water by measuring the circumference of the Earth with reasonable precision. Galileo showed that (b) and (c) are false statements, the first because the Earth circles the sun, the second because in a vacuum all bodies fall with the same acceleration. And Einstein's relativity theory predicted that time becomes exponentially slower as a body approaches the speed of light, falsifying statement (d). Bertrand Russell once pointed out that the theory of relativity showed that statement (d), like some other important commonsense beliefs, cannot withstand precise scientific examination (*Cf.* Russell 1925, Ch. 1; Popper 1972, Ch. 2, sec. 2). Finally, statement (e), affirming the seemingly commonsense corpuscular theory of light (defended by Newton, but already in some way evinced in Antiquity), has been found to be mistaken, since light consists of transverse waves (Huygens-Young theory), even though under certain conditions it behaves as though it consisted of particles (wave-particle theory).

The point I wish to emphasize, however, is that none of the five above-cited statements legitimately belongs to correctly understood common sense – a sense I call 'modest.' If we examine these statements more closely, we see they are in fact *extrapolations* grounded on statements of modest common sense. These extrapolations are of speculative interest and were made in the name of science by scientists and even by philosophers projecting ideas of common sense into new domains that would later belong to science. In my view, the true statements of common sense – the modest statements for which (a), (b), (c), (d) and (e) could be the corresponding non-modest extrapolations – are respectively the following:

(a') The Earth is flat.
(b') Each day the sun crosses the sky.
(c') Heavier bodies around us fall more rapidly than lighter ones.

(d') Time flows uniformly for all bodies around us, independently of their motion.
(e') Light has rays.

Now, what is at stake is that these statements have been made for thousands of years and have been confirmed thousands of times by everyday observation and continue to be confirmed, independently of the scientific development. It is obvious that (a') is a true statement if we understand it to mean that when we look at the world around us without having the ambition to generalize this observation to the whole Earth, we see that – discounting hills, valleys, and mountains – the landscape is obviously flat. Statement (b') is also true since it is anterior to the distinction between the real and the apparent motion of the sun. If we consider only the apparent motion of the sun, we see that the sentence 'The sun crosses the sky each day' can be considered true without implying that the sun revolves around the Earth. All it affirms is that in equatorial and sub-equatorial regions of the Earth we see that each day the sun rises in the East, crosses the sky, and sets in the West, which no sensible person would ever doubt.[1] Even after science proved that bodies of different masses fall with the same acceleration in a vacuum, statement (c') remains true for everyday experience. After all, it only affirms the commonplace notion that under ordinary conditions a light object such as a feather falls much more slowly than a heavy one such as a stone... Statement (d') also remains true, since it concerns the movements of things in our surroundings, leaving aside extremely high speeds or incredibly accurate measurements of time. (In everyday life, one would never need to measure time dilation, which is detectable only when a body approaches the speed of light and has nothing to do with everyday experience. In everyday life, no one ever comes home from a two-week bus trip to discover that family members are now many years older than before). Finally, (e') has been accepted, at least since Homer, as is shown by his poetic epithet 'rosy-fingered dawn.' And we often see sunbeams at dawn or dusk or peeping through gaps in the clouds on overcast days.

But then, what is the point in comparing statements (a)-(b)-(c)-(d)-(e) with the corresponding statements (a')-(b')-(c')-(d')-(e'), making the first set refutable by science, while the latter statements remain true? The answer is that scientifically or speculatively motivated commonsense

[1] This is a statement like that by Heraclitus of Ephesus, who noted that 'The sun is the width of a human foot.' As an interpreter wrote, we need only lie on the ground and hold up a foot against the sun to see that this is true.

statements exemplified by (a)-(b)-(c)-(d)-(e) have very often been viewed equivocally as if they were legitimate commonsense statements. However, statements of modest common sense like (a')-(b')-(c')-(d')-(e') are the only ones naturally originating from community life, being omnipresent in the most ordinary linguistic practices. They continue to be perfectly reliable despite the theoretical conclusions of Galileo and Einstein, since their truth is independent of science. The contrast between these two kinds of example shows how mistaken the claim is that many or most commonsense truths have been refuted by science. What science has refuted are extrapolations of commonsense truths by scientists and philosophers who have projected such humble commonsense truths beyond the narrow limits of their original context. If we take into account the aforementioned distinction, we find a lack of conflict between the discoveries of science and the claims of commonsense wisdom, also including ones used as examples by philosophers like G. E. Moore. [2]

I do not claim modest commonsense truths are in principle irrefutable, but only that no one has managed to refute them. Nothing warrants, for instance, asserting that from now on the world around us will be different in fundamental ways. A statement like (b') can be falsified. Perhaps for some unexpected reason, the Earth's rotation on its axis will slow down so much that the sun will cease its apparent movement across the sky. In this case, (b') would also be refuted for our future expectations. But even in this case, (b') remains true concerning the past, while the corresponding ambitious extrapolation (b) has always been false. In fact, all I want to show is that true commonsense statements – modest ones – are much more reliable than scientifically oriented minds believe, and science has been unable to refute them, insofar as we take them at their proper, humble face value.

Similar reasoning applies to the *a priori* knowledge of common sense. To justify this new claim, consider first the case of statements like (i) 'Goodness is praiseworthy,' which is grammatically identical with

[2] I am unable to find real exceptions. Under normal circumstances, fire has always burned. Some say that the idea that trees draw energy from the earth was once a commonsense truth until photosynthesis was discovered… But this idea wasn't a very basic or modest commonsense truth since it could easily be refuted by the well-known fact that trees do not grow in complete darkness. The idea that a *new* sun crosses the sky each new day is surely absurd – but is it a commonsense idea? In fact, it was suggested by a philosopher, Heracleitus, going beyond the humble intentions of modest common sense. Modest, humble common sense is not interested in answering such questions, which have no relationship to ordinary life concerns.

statements like (ii) 'Socrates is wise.' Both have the same superficial subject-predicate grammatical structure. Since in the first case the subject 'Goodness' does not designate any object accessible to the senses, Plato would have concluded that this subject must refer to 'goodness-in-itself': the purely intelligible *idea* of goodness, existing in an eternal and immutable non-visible realm only accessible to the intellect. Plato reached his conclusion based on the commonplace grammatical distinction between subject and predicate found in natural language. Under this assumption, he was likely to see a statement like (iii) 'Goodness in itself exists' as a commonsensical truth. In fact, according to his doctrine, it should be an *a priori* truth.

However, we know that with Frege's introduction of quantificational logic at the end of the 19th century, it became clear that statements like (i) should have a deep logical structure that is much more complex than the subject-predicate structure of (ii). Statement (i) should be analyzed as saying that all good things are praiseworthy, or (iv) 'For all *x*, if *x* is good, then *x* is praiseworthy,' where the supposed proper name 'Goodness' disappears and is replaced by the predicate '… is good.' This new kind of analysis reduced considerably the pressure to countenance the Platonic doctrine of ideas.

However, the suggestion that the subject 'Goodness' refers to an abstract idea clearly does not belong to modest common sense, and statement (iii), 'Goodness in itself exists,' isn't even inscribed in our natural language. It also belongs to ambitious common sense. Statement (iii) was a speculative extrapolation by a philosopher based on an implicit appeal to the superficial grammar of natural language, and although (iii) is probably false, it would be unjust to blame modest common sense and our ordinary language intuitions on subject-predicate grammar. Finally, it is wise to remember that quantificational truth-functional logic has not undermined the (commonsensical) grammar of our natural language; it has only selected and made us conscious of vastly extended fundamental patterns underlying the representative function of natural language.

What all these examples do is to undermine the frequently made claim that scientific progress contradicts common sense. Scientific discoveries only refute speculative extrapolations of common sense and natural language made by scientists and philosophers, such as the idea that the Sun revolves around the Earth or that there is a purely intelligible world made up of abstract ideas like that of Goodness in itself. But nothing of the sort has to do with the explanations given by modest common sense, the only ones long established by mankind's shared practical experience down through the ages.

3. Resisting changes in worldviews

Finally, I wish to consider commonsense ideas that are challenged by changes in our worldviews. This is, for instance, the case with the belief that a personal God exists or that we have minds independently of our bodies. The objection is the following. The overwhelming majority of cultures accept a God (or gods) and the soul as undeniably real. In Western Civilization, for the last two-thousand years society has even sanctioned denial of these beliefs with varying degrees of severity, sometimes even resorting to capital punishment. Although they were once commonsense beliefs, today no one would say that they are almost universally accepted. On the contrary, few scientifically educated persons would agree with them. Consequently, it seems that common sense ideas can change in response to changes in our worldviews...

My reaction to this does not differ very much from my response to the objection contrasting common sense with the progress of science. Beliefs regarding our worldviews lack universality, not really belonging to what could be called modest common sense. There are entire civilizations, particularly in Asia, where the idea of a personal God is foreign to the dominant religion. Regarding the soul, I remember a story told by an anthropologist who once asked a native Brazilian what happens after people die. The native answered: – 'They stay around.' – 'And later?' asked the anthropologist. – 'They go into some tree.' – 'And then?' – 'Then they disappear'...[3] The lack of concern was evident. And the unavoidable conclusion is that belief in a personal God and an eternal soul do not enjoy the kind of universality that would be expected of modest common sense; if they are said to belong to common sense, this must be an ambitious form of common sense. In fact, these beliefs seem to result from the distortion of ordinary views through *wishful thinking*, which has happened particularly in Western culture.[4]

Natural language also supports the view that these beliefs are not chiefly commonsensical: a person holding religious beliefs usually does not say he *knows* that he has a soul independent of his body… He prefers to claim he *believes* in these things. And even this belief has a particular

[3] Roberto DaMatta, in an interview. (A more forceful example is the obstinate rejection of any kind of theism of the Pirahã tribe in the Amazon rainforest studied by Daniel L. Everett).

[4] It was certainly much easier to believe in the existence of a personal God and an eternal soul independent of the body a thousand years ago, before the steady accumulation of divergent knowledge discovered by the natural and human sciences.

name: 'faith,' which is belief not supported by reason and observation (against faith there are no arguments). On the other hand, the same person would never deny that he *knows* there is an external world and that he *knows* this world existed long before he was born… Modest commonsense knowledge is not a question of wishful thinking or non-rational faith.

What all these arguments suggest is that *modestly understood commonsense truths – together with the very plausible discoveries of real science – can reasonably be said to form the basis of our rationality, the highest tribunal of reason.* Furthermore, since science itself can only be constructed starting from a foundation of accepted modest commonsense beliefs, it does not seem possible, even in principle, to deny modest common sense *as a whole* on the authority of science without also having to deny the very foundations of rationality.

Not only do science and changes in our worldview seem unable to refute modest common sense, even skeptical hypotheses cannot do this in the highly persuasive way one could expect. Suppose, for instance, that radical skeptics are right, and you discover that until now you have lived in what was just an illusory world… Even in this case, you would be unable to say that the world where you lived until now was *unreal* in the most important sense of the word. For that world would still be fully real in the sense that people perceived it with maximal intensity, and it was independent of the will, was interpersonally accessible and obeyed natural laws… These are criterial conditions that when together satisfied create our conventional sense of reality, a sense in itself indefeasible even by skeptical scenarios (See Ch. VI, sec. 29).

4. Primacy of Established Knowledge

The upshot of the comparison between modest common sense and science is that we can see science as not opposed to modest common sense, but rather as its proper extension, so that both can be mutually supportive. According to this view, *science is expanded common sense*. Contrary to Wilfrid Sellars (1962: 35-78), the so-called 'scientific image of the world' did not develop in opposition to or even independently of the old 'manifest image of the world,' for there is no conflict between them. This conclusion reinforces our confidence that underlying everything we can find commonsense truths, insofar as they are judiciously identified and understood.

In endorsing this view, I do not claim that unaided modest commonsense truth can resist philosophical arguments, as philosophers like Thomas Reid seem to have assumed. One cannot refute Berkeley's

anti-materialism by kicking a stone or answer Zeno's paradox of the impossibility of movement by putting one foot in front of the other. These skeptical arguments must be wrong, but to disprove them, philosophical arguments are needed to show *why* they seemingly make sense, again grounding their rejection at least partially in other domains of common sense if not science, something reached only by the comprehensiveness of philosophical reasoning. Hence, what I wish to maintain is that the principles of modest common sense serve as the most reliable assumptions and that some fundamental modest commonsense principles will always be needed if we do not wish to lose our footing in everyday reality.

I reject the proposal that a philosophy based on modest common sense and its effects on natural language intuitions would be sufficient. It is imperative to develop philosophical views compatible with and complementing modern science. We must construct philosophy on a foundation of common sense *informed* by science. That is: insofar as formal reasoning (logic, mathematics…) and empirical science (physics, biology, psychology, sociology, neuroscience, linguistics...) can add new extensions and elements beyond modest commonsense principles, and these extensions and elements are relevant to philosophy, they should be taken into account. As we saw above, it was through the findings of predicate calculus that we came to know that the subject 'goodness' in the sentence 'Goodness is praiseworthy' should not be logically interpreted as a subject referring to a Platonic idea, since what this sentence really means is 'For all *x*, if *x* is good, *x* is praiseworthy.'

I will use the term *established knowledge* for the totality that includes modest commonsense knowledge and all the extensions the scientific community accepts as scientific knowledge. Any sufficiently well-informed and reasonable person would agree with this kind of knowledge, insofar as he would be able to properly understand and evaluate it. It is in this revised sense that we should reinterpret the Heraclitean *dictum* that we must rely on common knowledge as a city relies on its walls.

The upshot of these methodological remarks is that we should judge the plausibility of our philosophical ideas against the background of established knowledge, that is, comparing them with the results of scientifically informed common sense. We may call this the principle of the *primacy of established knowledge*, admonishing us to make our philosophical theses consistent with it. Philosophical activity, particularly as descriptive metaphysics,[5] should seek *reflexive equilibrium* with the

[5] The expression 'descriptive metaphysics' was introduced by P. F. Strawson in contrast to 'revisionary metaphysics.' It aims to describe the most general features

widest possible range of established knowledge, the knowledge mutually supported by both modest common sense and scientific results. This is the ultimate source of philosophical credibility.

Finally, if we find inconsistencies in challenging speculative philosophical theories because they seem to debunk much or our established knowledge, we should treat them as paradoxes of thought, even if they can be very instructive, and should search for arguments that reconcile these products of philosophical speculation with established knowledge. Lacking reconciliation, we should treat philosophical theses only as *proposals*, even if they can be sometimes extraordinarily stimulating from a speculative viewpoint, as is the case of revisionary metaphysics, superbly exemplified by Leibniz, Berkeley and Hume and in some measure also by most American analytic philosophers since W. V-O. Quine. This acknowledgment does not mean that their results require acceptance as 'solid' discoveries, but rather that they deserve attentive consideration, the sort we grant to the best cases of expansionist scientism. To proceed otherwise can lead us down the slippery slope to dogmatism.

5. Philosophizing by examples

We must complement our methodological principle of the primacy of established knowledge with what Avrum Stroll called the *method of philosophizing by examples* (1998, x-xi). He himself used this method to construct relevant arguments against Putnam's externalism of meaning.

Stroll was a Wittgenstein specialist, and Wittgenstein's therapeutic conception of philosophy directly inspired his approach. According to Wittgenstein, at least one relevant way of doing philosophy is by performing philosophical therapy. This therapy consists in comparing the speculative use of expressions in philosophy – which is very often misleading – with a variety of examples, most of them of their everyday usage – where these expressions earn their proper meanings, using a method of similarity and contrast to clear up the confusion. He thought this therapy was only possible through meticulous comparative examination of various real and imaginary concrete examples of intuitively correct and even incorrect uses of expressions. This would make it possible to clarify the true meanings of our words so that the hidden absurdities of

---

of our actual conceptual schema, while revisionary metaphysics attempts to provide a new schema to understand the world. Strawson, Aristotle, and Kant developed descriptive metaphysics, while Leibniz and Berkeley developed revisionary metaphysics (Strawson 1991: 9-10).

metaphysics would become evident... Since contemporary philosophy of language tends to be unduly metaphysically oriented, and in this way diametrically opposed to the kind of philosophy practiced by Wittgenstein, a similar critique of language, complemented by theoretical reflection, is what much of contemporary philosophy needs in order to find its way back to truth.

I intend to show that today's metaphysics of reference and meaning suffers from a failure to consider adequately, above all the subtle nuances of linguistic praxis. It suffers from an accumulation of potentially obscurantist products of what Wittgenstein called 'conceptual houses of cards' resulting from 'knots of thought' – subtle semantic equivocations caused by a pressing desire for innovation combined with a lack of more careful attention to nuanced distinctions of meaning that expressions receive in different contexts where they are profitably used, also because the reason why they might be interesting is that they are by queer ways magnifying some real insight.

One criticism of Wittgenstein's therapeutic view of philosophy is that it would confine philosophy to the limits of the commonplace. Admittedly, there is no good reason to deny that the value of philosophy resides largely in its theoretical and systematic dimensions, in its persistent attempt to make substantive generalizations. I tend to agree with this, since I also believe that in its proper way philosophy can and should be theoretical, even speculatively theoretical. Nonetheless, I think we can to a great extent successfully counter this objection to Wittgenstein's views, first interpretatively and then systematically.

From the interpretative side, we have reason to think that the objection misunderstands some subtleties of Wittgenstein's position. The most authoritative interpreters of Wittgenstein, G. P. Baker and P. M. S. Hacker, insisted that he did not reject philosophical theorizing *tout court.* In rejecting philosophical theorizing, he was opposing *scientism*: the kind of philosophical theorizing that mimics science. Scientism tries to reduce philosophy itself to science in its procedures, range, and contents, as he personally saw happening in logical positivism.[6] Instead, he would

[6] As these interpreters wrote: 'Wittgenstein's objection to "theorizing" in philosophy is an objection to assimilating philosophy, whether in method or product, to a theoretical (super-physical) science. But if thoroughgoing refutation of idealism, solipsism or behaviorism involves a theoretical endeavor, Wittgenstein engages in it.' (Baker & Hacker 1980: 489) Anthony Kenny (1986) preferred to think that Wittgenstein actually held two competing views on the nature of philosophy – therapeutic and theoretical. But the here proposed unified interpretation seems more charitable.

countenance a different sort of theorizing, particularly the 'dynamic,'[7] the 'organic' instead of 'architectonic' (Wittgenstein 2001: 43) – a distinction he seems to have learned from Schopenhauer (Hilmy 1987: 208-9). This helps explain why, in a famous passage of *Philosophical Investigations*, he argued that it is both possible and even necessary to construct *surveillable representations* (*übersichtliche Darstellungen*). These can show the complex logical-grammatical structure of the concepts making up the most central domains of understanding. As he wrote:

> A main source of our failure to understand is that we do not command a clear view of the use of our words – Our grammar is lacking in this sort of surveillability. A surveillable representation produces just that understanding which consists in 'seeing connections'; hence the importance of finding and inventing intermediate cases. The concept of surveillable representation is of fundamental significance for us. It earmarks the form of account we give, the way we look at things (Is this a '*Weltanschauung*'?). (1984c, sec. 122)

Now, in a sense, a surveillable representation must be theoretical since it must contain *generalization,* and this constitutes the ultimate core of what the word 'they' means. (Well aware of this, Karl Popper famously called the statement 'All swans are white' a *theory*, adding that this theory was falsified by the discovery of black swans in Australia…) If we agree that all generalizations are theoretical, any surveillable representation, as it must contain generalizations, must also be theoretical.

Moreover, the addition of *intermediate connections* already existent but not explicitly named by the expressions of ordinary language justifies our making explicit of previous well-grounded conventions that serve as links connecting a multitude of cases. It is possible that because of the generality and function of these links, they never need to emerge in linguistically expressible forms (consider, for instance, our MD-rule for proper names). Expositions of these links are properly called 'descriptive', insofar as they are already present under the surface of language. But it is fully acceptable to call them 'theoretical' – in the sense of a description of general principles inherent to natural language – if they are intended to be the right way to assure the unity in diversity that our usage of expressions is able to achieve.

The addition of intermediary connections helps to explain why normal language philosophy, as initially developed by Gilbert Ryle and J. L.

[7] As he writes, 'We have now a theory, a "dynamic" theory (Freud speaks of a "dynamic" theory of dreams) of the sentence, of the language, but it appears to us not as a theory.' (*Zettel* 1983b: 444).

Austin gradually transformed itself into far more liberal and theoretical forms of philosophy inspired by natural language that we can already find in some works of P. F. Strawson and later in H. P. Grice[8] and John Searle. It also helps to justify the introduction of new *technical terms* to fill the gaps in natural language. Terms like 'criterion,' 'language-game,' 'grammatical sentence,' 'forms of life' and even 'surveillable representation' support this point in Wittgenstein's own writings. In fact, even Austin, the chief defender of a quasi-lexicographical ordinary language analysis didn't eschew the creation of new technical terms. Expressions like 'locutionary act' (composed of 'phonetic', 'phatic' and 'rhetic acts'), 'illocutionary act' and 'perlocutionary act' (1962, Lect. VIII) were created as the only way to express – guided by reasoning on interactive linguistic activity – fundamental deep structures totally unexpressed in our normal usage.

Now, from the systematic argumentative side, we can say that independently of the way we interpret Wittgenstein, there are good reasons to believe theoretical considerations are indispensable. An important point is that philosophy can only be therapeutic or critical because its work is inevitably based on theoretical, that is, generalized assumptions that make possible its therapeutic efficacy. Usually, Wittgenstein did not explicitly state or develop the assumptions needed to make his conceptual therapy convincing. He was an intuitive thinker in the style of Heracleitus or Nietzsche. Because of this, he all too often did not develop his insights beyond the epigrammatic level. In any case, general assumptions are

[8] Paul Grice's sophisticated and ingenious work contains an influential (albeit qualified) criticism of ordinary language philosophy as practiced by Ryle, Austin and Strawson (1989, Chs. 1, 2, 10, 15, 17). According to him, these philosophers often confused ordinary uses of statements resulting from conversational implicatures with their literal meaning. When implicature failed, they mistakenly concluded that these statements had no meaning. This would be the case of statements like 'This flag looks like red' (supposedly understood by Austin as showing that sense-data do not exist because this statement is devoid of sense), 'The present King of France is wise' (understood by Strawson as a statement without truth-value) and 'If green is yellow then 2 + 2 = 5' (understood by him as showing the odd character of material implication). I agree with Grice's rejection of all these ordinary language philosophers' conclusions, even if I remain suspicious regarding his own explanations. Material implication, for instance, still belongs to our practice of truth-functional reasoning, which makes explicit a basic general layer subsumed under our more informative factual language. In this sense, it also provides wide intermediate connections. That is, under sufficiently critical scrutiny, natural language intuitions still provide a valuable guide – a point with which Grice would certainly agree.

inevitable if our aim is expose equivocal views in an efficacious way: The critical (therapeutic) and the more constructive (theoretical) searches for surveillable representations can be understood as two complementary sides of the same analytical coin (Costa 1990: 7 f.). Theoretical assumptions are the indispensable active principle of logic-conceptual therapeutic potions.

Recapitulating, we have found two main methodological principles for orienting our research in this book:

A. The principle of the *primacy of established knowledge* (our principle of all principles), according to which modest common sense, complemented by scientific knowledge, constitutes the highest tribunal of reason in judging the plausibility of philosophical views.
B. The method of *philosophizing by examples*, according to which the best way to orient ourselves in the philosophical jungle is to test our ideas in all possible opportunities by analyzing a sufficient number of different examples. If we do not use this method, we risk losing ourselves in a labyrinth of empty if not fallacious abstractions.

Oriented by the two above-considered methodological principles, I intend to accomplish several tasks, which are fundamental if we wish to put philosophy again on the right track.

6. Tacit knowledge of meaning: traditional explanation

I will assume the practically indisputable notion that language is a system of signs basically governed by conventionally grounded rules, including semantic ones. Linguistic conventions are rules obeyed by most participants in the linguistic community. These participants expect other participants to comply with similar or complementary rules and vice-versa, even if they aren't really aware of them (*Cf.* Grice 1989, Ch. 2; Lewis 2002: 42). According to this view, the sufficiently shared character of language conventions is what makes possible the use of language to communicate thoughts.

One of the most fundamental assumptions of the old orthodoxy in philosophy of language is that we *lack awareness* of the effective structures of semantically relevant rules governing the uses of our language's most central conceptual expressions. We know how to apply the rules, but the rules are not available for explicit examination. Thus, we are unable to command a clear view of the complex network of tacit agreements

involved. The reason is the way we learn expressions in our language. Wittgenstein noted that we learn the meaning-rules governing the correct use of our linguistic expressions not by means of explicit definitions, but by *training* (*Abrichtung*), that is, through informal practice, imitation, and correction by others who already know how to use them properly. Later analytic philosophers, from Gilbert Ryle to P. F. Strawson, Michael Dummett, and Ernst Tugendhat, have always insisted that we do not learn the semantically relevant conventions of our language (i.e., the semantic-cognitive rules determining referential use of expressions) through verbal definitions, but rather in non-reflexive, unconscious ways. Tugendhat wrote that we learn many of these rules in childhood through ostension by means of positive and negative examples given in interpersonal contexts: other speakers confirm them when correct and disconfirm them when incorrect. Hence, the final proof that we understand these rules is interpersonal confirmation of their correct application. (Tugendhat & Wolf 1983: 140) For this reason, it is often so hard or seemingly impossible to obtain an explicit verbal analysis of the meaning of an expression that is really able to reveal its meaning-rules. Using Gilbert Ryle's terms, with regard to these meaning-rules we have *knowing how*, i.e., skill, competence, an automatized ability that enables us to apply them correctly; but this is insufficient to warrant *knowing that*, namely, the capacity to report what we mean verbally (1990: 28 f.).

This non-reflexive learning of semantic rules applies particularly to philosophical terms like 'knowledge,' 'consciousness,' 'understanding,' 'perception,' 'causality,' 'action,' 'free will,' 'goodness,' 'justice,' 'beauty,' which are central to our understanding of the world (Tugendhat 1992: 268). Because of their more complex conceptual structure and internal relationships with other central concepts, these concepts are particularly elusive and resistant to analysis, opening room to the most various intentions. This insight certainly also applies to conceptual words from philosophy of language, like 'meaning,' 'reference,' 'existence' and 'truth,' which will be examined later in this book. Finally, to make things more complicated, relevant concepts are also in a sense empirically grounded and not completely immune to additions and changes resulting from the growth of our knowledge. For instance: before recent advances in neuroscience, bodily movement was considered essential to the philosophical analysis of the concept of action. Now, with sensitive devices able to respond to electrical discharges in our motor-cortex, we are able to move external objects using sheer willpower. Intentions unaided by bodily movements are now sufficient to produce external physical motions intended by the agent (See *neuroprosthetics* and BCIs).

However, lack of semantic awareness can become a reason for serious intellectual confusion when philosophers try to explain what these terms *mean*. Philosophers are very often under the pressure of some generalizing purpose extrinsic to that required by the proper nature of their object of investigation. Consider theistic purposes in the Middle Ages and scientistic purposes in our time, which can easily produce startling but erroneous magnifications hinging on minor real findings. Wittgenstein repeatedly expressed these metaphilosophical views throughout his entire career. Here are some of his best quotes, in chronological order, beginning with his *Tractatus Logico-Philosophicus* and ending with his *Philosophical Investigations*:

> Natural language is part of the human organism and not less complicated than it. ... The conventions that are implicit for the understanding of natural language are enormously complicated. (1984g, sec. 4.002)

> Philosophers constantly see the method of science before their eyes and are irresistibly tempted to ask and answer questions the way science does. This tendency is the real source of metaphysics, and leads the philosopher into complete darkness. (1958: 24)

> We can solve the problems not by giving new information, but by arranging what we have always known. Philosophy is a battle against the bewitchment of our intellect by language. (1984c sec. 109)

> The aspects of things that are most important for us are hidden because of their simplicity and familiarity. (One is unable to notice something – because it is always before one's eyes.) The real foundations of his inquiry do not strike a person at all. Unless that fact has at some time struck him. – And this means: we fail to be struck by what, once seen, is most striking and most powerful. (1984c, sec.129)

> Contrary to empirical statements, rules of grammar describe how we use words in order to both justify and criticize our particular utterances. But as opposed to grammar book rules, they are not idealized as an external system to be conformed to. Moreover, they are not appealed to explicitly in any formulation, but are used in cases of philosophical perplexity to clarify where language misleads us into false illusions … (A whole cloud of philosophy is condensed into a drop of grammar.) (1984c, II xi)

Around the mid-twentieth century, a number of analytical philosophers were in significant ways directly or indirectly influenced by Wittgenstein views. They believed clarification resulting from the work of making explicit the tacit conventions that give meaning to our natural language

was a kind of revolutionary procedure: We should identify most if not all philosophical problems with conceptual problems that could be solved (or dissolved) by means of conceptual analysis.

Notwithstanding, except for the acquisition of new formal analytical instruments and a new pragmatic concern leading to more rigorous and systematic attention to the subtleties of linguistic interaction, there was nothing truly revolutionary in the philosophy of linguistic analysis and the critique of language associated with it. Analysis of the meaning of philosophically relevant terms as an attempt to describe the real structure of our thinking about the world is no more than the resumption of a project centrally present in the whole history of Occidental philosophy. Augustine wrote: 'What, then, is time? If no one asks me, I know; if I wish to explain it to him who asks, I know not.' (2008, lib. XI, Ch. XIV, sec. 17) In fact, we find the same concern already voiced by Plato. If we examine questions posed in Plato's Socratic dialogues, they all have the form 'What is *X*?', where *X* takes the place of philosophically relevant conceptual words like 'temperance,' 'justice,' 'virtue,' 'love,' 'knowledge'… What then follows are attempts to find a definition able to resist objections and counterexamples. After some real progress, discussion usually ends in an aporetic way due to merciless conceptual criticism. That is, philosophy based on analysis of conceptual meaning has always been with us. It is the main foundation of our philosophical tradition, even when it is hidden behind its most systematic and speculative forms.[9]

Finally, by defending the view that philosophy's main job is to analyze implicit conceptual knowledge, I am not claiming that philosophy cannot be about the world, as some have objected (Magee 1999, Ch. 23). Even as an inquiry turned to our conceptual network, philosophy continues to be about the world, because the concepts analyzed by philosophy are in one way or another about the world. Moreover, in a systematic philosophical work, central concepts of our understanding of the world are analyzed in their internal relations with other central concepts, with the same result that philosophy is indirectly also about the world – about the world as it is synthetically reflected by the central core of our conceptual network.[10]

---

[9] Philosophers like Berkeley, Leibniz and Hegel can be seen as doing revisionary conceptual *analysis*, refuting and replacing ambitious interpretations of common sense by new ones.

[10] Rudolf Carnap's *formal mode of speech* (1937, part 5, sec. A, § 79) instead of *material mode of speech*, and particularly W. V-O. Quine's broader *semantic ascent* (1960, Ch. 7, § 56) point to this same fact, namely, that by means of conceptual analysis we emphasize linguistic forms only in order to have a clearer sight of them.

Indeed, even if the philosophical analysis of our conceptual structures does not depend on empirical experience as such, empirical experience has already in one way or another entered into the production and change of such conceptual structures.

7. A very simple model of a semantic-cognitive rule

We urgently need to clarify the structures of our semantic-cognitive rules as the concept is used here. However, it is not very helpful if we begin by attempting to analyze a conceptual rule constitutive of a philosophical concept-word. Not only are these concept-words usually polysemic, but the structures of central meaning-rules expressed by them are much more complex and harder to analyze and in this way to characterize or define. Anyway, although philosophical definitions can be extremely difficult to achieve, the skeptical conclusion that they are impossible can well be too hasty.

To get a glimpse into the nature of a semantic-cognitive rule – in the case an ascription rule[11] – I strategically chose a very trivial concept-word, since its logical grammar is correspondingly easier to grasp. Thus, I wish to scrutinize here the standard meaning of the concept-word 'chair,' using it as a simple model that can illustrate our approach to investigating the much more complicated philosophical concepts. We all know the meaning of the word 'chair,' though it would not be so easy to give a precise definition if someone asked for one. Now, following Wittgenstein's motto, according to which 'the meaning of a word is what the explanation of its meaning explains' (1984g, sec. 32), I offer a very reasonable definition (explanation) of the meaning of the word 'chair.' You can even find something not far from it in the best dictionaries. This definition expresses the characterizing ascription rule of this concept-word, which is the following:

(C) *Chair (Df.) = a non-vehicular seat with a backrest, designed for use by only one person at a time* (*it usually has four legs, sometimes has armrests, is sometimes upholstered, etc.*).[12]

[11] A good glimpse into the nature of identification rules was already provided in the Appendix of Chapter I.

[12] If you wish to avoid the word 'seat', you can also define a chair as 'a moveable piece of furniture with a raised surface and a backrest, made for only one person at a time to sit on.'

In this definition, the conditions stated outside of parentheses are *necessary* and together *sufficient*: first a chair must be a non-vehicular seat (since seats in cars and airplanes… aren't called chairs); second, a chair must be a seat with a backrest (since without a backrest it would be only a stool, a sattle seat, etc.); and third, it must be an artifact designed for a single person to seat at a time. These criterial conditions form an *essential*, indispensable condition, also called the *definitional or primary criterion* for the applicability of the concept-word, to use Wittgenstein's terminology.

What follows in parentheses are complementary (dispensable) *secondary criteria* or *symptoms:* usually, a chair has four legs, often it has armrests, and sometimes it is upholstered. These indications can be helpful in identifying chairs, even though they are irrelevant if the definitional criterion isn't satisfied. A chair need not have armrests, but there cannot be a chair with armrests but no backrest (this would be a bench). Thus, with (C) we have an expression of our implicit conventional ascription rule for the general term 'chair,' which should belong to the domain of what Frege calls *sense* (*Sinn*).[13]

Though I do not think that this definition couldn't be improved (or changed), I find it hard to oppose it. Table-chairs, armchairs, easy chairs, rocking chairs, wheelchairs, beach chairs, electric chairs, thrones… all conform to the definition (a kneeling chair without backrest is rather a *so-called* chair). The definition gives the core of the conceptual meaning, while the different types of chair are recognized as sub-concepts adding to the core a diversity of sub-conceptual modulations.[14] However, car, bus

[13] As will be frequently recalled, I do not deny that referential meanings include things that cannot be really captured by descriptive conventions, unlike case (C) – things like perceptual images, memory-images, feelings, smells. However, they belong much more to the semantic level called by Frege *illuminations* (*Beleuchtungen*), based on *natural regularities* more than on conventions.

[14] You can continue using the word 'chair' when pointing to a wheelchair, but even this semantic flexibility is already definitionally sustained, insofar as you are allowed to identify a chair with wheels as a chair. The sub-conceptual semantic modulation here is nothing but an addition to the definition of a chair. There are well-known cases like (i) 'cut' (*Df.*): using a sharp-edged or another device to separate something into parts, (ii) 'love' (*Df.*): an intense feeling of affection, (iii) 'game' (*Df.*): an activity with rules intended to be used as a means of entertainment (*pace* Wittgenstein), and (iv) 'abuse' (*Df.*): damaging use of something. These definitions, even if incomplete, express meaning-giving conventional cores that can receive conventional sub-conceptual supplements by the addition of words in expressions such as for (i) 'cut the cake,' 'cut the grass,' for (ii) 'love a woman,' 'love a child' ('loving chocolate' is already an extended, metaphorical use), for

and airplane seats are not called 'chairs' but by convention seats, only because they are made to bring persons in vehicles from a place to another. Anyway, we can always imagine *borderline* cases. There could be a seat whose backrest is only 20 cm. high (is it a stool or a chair?), a chair with a seat raised only 10 cm. above the floor (is it even a seat?), a chair whose backrest was removed for some hours (did it become a backless chair or provisionally a stool?). Suppose we find a tree trunk in a forest with the form of a chair that, with some minor carving and painting, is now being used as a chair (it was not manufactured as a chair, but minor changes turned it into something we could maybe call a real chair, depending on the relevance of the changes). Nevertheless, our definition is still reasonable despite vague borderline cases. Empirical concepts all have some degree of vagueness, and one can even argue that vagueness is a metaphysical property of reality. Indeed, if our definition of a chair had overly sharp boundaries, it would be inadequate, since it would not reflect the desired flexibility of application belonging to our normal word 'chair,' tending to stiffen the extension of the concept.

An often overlooked point is that language is a pragmatic device: *what really justifies a semantic-cognitive rule is its practical applicability to common cases*. That is, what really matters are cases to which we can apply the ascription rule without much hesitation and not those rare borderline cases where we cannot know if the ascription rule is definitely applicable, since the rarity of these cases, so much praised by philosophers, makes them irrelevant from a practical point of view. Accordingly, the function of a concept-word is far from being discredited by a few borderline cases where we are at a loss to decide whether it is still applicable.

Furthermore, we need to distinguish real chairs from 'so-called chairs,' because in such cases we are making an *extended* or even a *metaphorical* use of the word. A child's toy chair, like a sculptured chair, is a chair in an extended sense of the word. In Victor Hugo's novel *Toilers of the Sea*, the main character ends his life by sitting on a 'chair of rock' on the seashore, waiting to be swept away by the tide... But it is clear from our definition that this use of the word is metaphorical: a real chair must be *made* by someone since it is an artifact, but the immovable stone chair was only a

---

(iii) 'play chess,' 'play tennis,' 'play solitaire,' and for (iv) 'abuse a drug,' 'abuse a child.' Such sub-conceptual modulations can also be made without a subsidiary word, by the context alone. Contemporary philosophy has in my view an insufficiently justified *bias* against definitions. (For a somewhat different view, see Recanati 2010: 29 f.)

natural object accidentally shaped by erosion into the rough form of a chair and then used as a chair.

There are also cases that only seem to contradict the definition, but that on closer examination do not. Consider the following two cases, already presented as supposed counterexamples (Elbourne 2011, Ch. 1). The first is the case of a possible world where some people are extremely obese and sedentary. They require chairs that on the Earth would be wide enough to accommodate two or three average persons. Are they benches? The relevant difference between a bench and a chair is that chairs are artifacts made for only one person to sit on, while benches are made wide enough for more than one person to sit on at a time. Hence, in this possible world what for us look like benches are in fact chairs, since they are constructed for only one sitter at a time. If these chairs were 'beamed' over to our world, we would say that they remained chairs, since the makers intended them to be chairs, even if we used them as benches. The second counterexample is that of a social club with a rule that only one person at a time can use each bench in its garden. In this case, we would say they continue to be benches and not chairs, since they are still artifacts designed for more than one person to sit on, even if they are now limited to single sitters. Elbourne also asked if a chair must have four legs. The answer is obvious since according to our definition having four legs isn't a defining feature: there are chairs with no legs, like an armchair, chairs with three legs, and we can imagine a chair with a thousand legs. The property of having four legs is what we have called a symptom or a secondary criterion of 'chair-ness,' only implying that a randomly chosen chair will probably have four legs.

One can always imagine new and more problematic cases that do not seem to fit the definition, but if we look at the definition more closely we discover that the difficulty is only apparent or that these 'exceptions' are borderline cases or that they are extensions or metaphors. Perhaps the definition indeed deserves some refinement, remembering that refinement isn't a mere change to something else.

Finally, the boundaries of what we call a 'chair' can also undergo changes from language to language and over time; in French an armchair (easy chair) is called a 'fauteuil' in contrast to a 'chaise' (chair), though a French speaker would agree that it is a kind of chair. I suspect that thousands of years ago, in most societies one could not linguistically distinguish a stool from a chair since a seat with a backrest was a rare piece of furniture until some centuries ago. Finally, the conventional similarities are here much more impressive than the dissonances, and these similarities seem to increase with the centrality of our concepts.

8. Criteria *versus* symptoms

To make things clearer, it is already worthwhile to broaden our consideration of Wittgenstein's distinction between *criteria* and *symptoms*. A *symptom* or a *secondary criterion* is an entity *E* that – assuming it is really given – only makes our cognitive awareness *A* of *E more or less probable*. In contrast, a *definitional or primary criterion* is an entity *E* (usually appearing as a complex criterial configuration) that – assuming it is really given – makes our cognitive awareness *A* of *E beyond reasonable doubt* (Wittgenstein 1958: 24; 2001: 28).[15]

For instance, if I can see four chair legs under a table, this is a symptom of a chair, since it greatly increases the probability that a chair is behind the table. But if I perceive that what is visually given to me is 'a moveable seat with a backrest made for only one person to sit on,' this puts my cognitive awareness of a chair beyond doubt. The definition (C) expresses a *definitional criterion*, understood as such because its assumed satisfaction leaves no possibility to doubt that we can apply the ascription rule for the concept-word 'chair.'

I cannot guarantee with absolute certainty that entity *E* (criterion or symptom) is 'really given' because I accept that the products of human experience are inevitably fallible. Nonetheless, using this 'assumed given-ness' based on experience and an adequate informational background, we can establish a subjective degree of *probability* when a symptom is satisfied and a *practical certainty* when a criterion is satisfied. In this last case, we might claim there is a probability so close to 1 that we can ignore the possibility of error in the cognitive awareness *A* that entity *E* is given. (Correspondingly, one could also speak in this sense of a conditional necessity.)

Symptoms or secondary criteria can help us identify entity *E* using cognitive awareness *A*, even if we cannot regard *E* as necessary. However, symptoms are of no use unless definitional criteria are also met. Four legs and armrests that do not belong to a chair would never make a chair.[16]

---

[15] The correct interpretation of this distinction is a controversial issue that does not concern us here; I give what seems to me the most plausible and useful version.

[16] At first view, it seems that these logico-conceptual remarks appeal to old-fashioned semantic *definitions* leading us to the rejection of findings of modern empirical psychology (*Cf.* E. Margolis & S. Laurence, 1999, Ch. 1). But this is only appearance. Consider Eleanor Rosch's important results. She has shown that we are able to categorize under a concept-word much more easily and quickly by appealing to *prototypical cases* (*Cf.* Rosch, 1999: 189-206). For example, we can more easily recognize a sparrow as a bird than an ostrich or a penguin. In the same

Terms like 'criteria' and 'symptoms,' as much as 'conditions' have so-called *process-product ambiguity*. We can see them as (a) *dependent* (internal) elements of the rule that identifies what is given, but we can also see them as (b) something *independent* (external) satisfying the rule, which is really given in the world. Our semantic-cognitive rules are also criterial rules, able with the help of imagination to generate criterial configurations belonging to them internally as (a). Hence, we could say that definition (C) is the expression of a semantic-criterial rule with the form: 'If we accept that *E* is really given, we must conclude *A*,' where the conclusion *A* is our *awareness* with practical certainty that *E* is given.

One problem here is to know what this awareness means. My suggestion will be that we can equate this cognitive awareness with our acceptance of the existence and applicability of a network of external inferential relations once a semantic-cognitive rule is satisfied. The conceptual meaning of 'chair,' for instance, consists of internal relations expressed by a definitional rule (C). But our awareness of the application of this conceptual meaning arises as a maze of external relations resulting from the satisfaction of (C). For example, if I am aware that a chair exists, I can infer that it has a particular location, that I can sit on it or ask someone to sit on it, that I could possibly damage it, borrow it, loan it, etc. I can do this even if I have no real consciousness of the structure of the rule I applied to identify the chair.

9. Challenges to the traditional explanation (i): John McDowell

Supporters of semantic externalism have challenged the idea that the meanings of expressions consist in our implicit knowledge of their constitutive rules or conventions. According to their view, the meanings of

---

way, an ordinary chair with four legs can be recognized as a chair more easily and quickly than can a wheelchair or a throne. However, this does not conflict with our definition, since for us the psychological mechanism of recognition responsible for the performance is not in question, but rather the leading structure subjacent to it. We can often appeal to symptoms as the most usual ways to identify things. For instance, we identify human beings first by their faces and penguins first by their appearance, even if human faces and a penguin's appearance are only symptoms of what will be then confirmed by expected behavior, memories, genetic makeup, etc. Hence, the ultimate criterion remains dependent on a definition. (In one wildlife film fake penguins outfitted with cameras deceived real penguins. The trouble with these moronic birds is that they are overly dependent on innate, instinctive principles of categorization.)

expressions are predominantly related to physical and social-behavioral worlds, depending in this way only on objects of reference and supposedly also on neurobiological processes involving autonomous causal mechanisms. From this perspective, there is little room for discussing the conventionality of meaning.

As evidence for the externalist view, we can adduce our lack of awareness of the structure of semantic rules determining the linguistic uses of our words. If we lack awareness of senses or meanings, it might be that they could as meanings be instantiated to a greater or lesser extent in a non-psychological domain. If this is so, in principle cognitive (also called pre-cognitive) participation in meaning could be unnecessary. Meaning could result solely from autonomous causal mechanisms, not recoverable by consciousness. In opposition to Michael Dummett's 'rich' view of implicit meaning, John McDowell illustrated the externalist position on the referential mechanism of proper names, observing that:

> We can have the ability to say that a seen object is the owner of a familiar name without having any idea of how we recognize it. The assumed mechanisms of recognizing can be neural machinery [and not psychological machinery] – and its operations totally unknown to whoever possesses them. (2001: 178)[17]

Some pages later, McDowell (following Kripke) asserts that the referential function of proper names would not be explained by implicit conventionally based identification rules that can be descriptively recovered, because:

> The opinions of speakers on their divergent evidential susceptibilities regarding names are products of self-observation, as much as this is accessible, from an *external* point of view. They are not intimations coming from the interior, from a normative theory implicitly known, a receipt for the correct discourse which guides the behaviour of the competent linguist. (2001: 190)

This view is in direct opposition to the one I defend in this book, not as much because it can never be justified, but because it isn't the usual case. In what follows, I intend to show that usually the implicit application of internal semantic-cognitive rules based on criteria is indispensable for the referential function. Moreover, we have already begun to see that to have a reference, a usually tacit and unconscious cognitive element must be

[17] The expression in brackets appears in the author's footnote on this passage. In Dummett's more orthodox position, McDowell sees a relapse into the psychologism justifiably rejected by Frege.

associated with our expressions and should be instantiated at least in some measure and at some moment in the language user's head. For in no case is this clearer than with McDowell's main focus: proper names (*Cf.* Appendix of Chapter I).

Here is how we could argue against McDowell's view. If he were correct, an opinion about the given criterial evidence for the application of a proper name found through external observation of our referring behavior should be gradually reinforced by the cumulative consideration of new examples, that is, *inductively*. Even repetition of the same example would be inductively reinforcing! However, this is far from the case. Consider our characterizing semantic-cognitive rule (C) for applying the concept-word 'chair.' We can see from the start that (C) *seems* correct. We naturally *tend* to agree with (C), even if we have never considered any examples of the word's application. And this shows that speakers are indeed only confirming a recipe for the correct application that comes from inside, as a matter of tacit agreement among speakers… Admittedly, *after* we hear this definition, we can test it. Thus, we can imagine a chair without a backrest but see that it is really a stool, which isn't properly a chair. If we try to imagine a chair designed so that more than one person can sit on it, we will conclude that we should call it a sofa or a garden bench... We can understand supposed counterexamples only as means to confirm and possibly correct or improve the definition, thereby discovering its extensional adequacy in a non-inductive way. This specification of meaning seems to be simply a contemporary formulation of something Plato identified as *reminiscence* (*anamnesis*): the recalling of ideas to mind. We do not need to go beyond this, imagining all sorts of chairs (rocking chairs, armchairs, wheelchairs…) in order to reinforce our belief in the basic correctness of our intuitive definition. But we can do it in order to recognize its more precise delimitation.

Now consider the same issue from McDowell's perspective. Suppose he is right and our knowledge of the meaning of a common name like 'chair' were the result of self-observation from an *external* viewpoint. We could surely acquire more certainty that chairs are seats with backrests made for one person to sit on by observing the similarities among real chairs that we can see, remember or imagine. Inductively, the results would then be increasingly reinforced by agreement among observers about an increasing number of examples. As we already noted, even examples of people reaching a shared agreement by singling out thousands of identical classroom chairs should enable us to increase our conviction that we have the factually true evidential conditions for applying the concept-word 'chair.' But this makes no sense. Moreover, it is clear that

one does not need much reflection to recognize that the idea is absurd of definition (C) capturing a neuronal mechanism and not resulting from an implicit shared agreement. Furthermore, alone the explanation of the implicitly conventional identification rule for the proper name Aristotle investigated in the Appendix of the last chapter is sufficient to make this whole discussion idle.

We conclude, therefore, that the ascription rule made explicit in definition (C) does, in fact, have the function of rescuing for consciousness the tacit convention governing the referential use of the word 'chair' (as with our earlier definition of 'Aristotle' in the Appendix of Chapter I). It seems from the start intuitive and may only require the help of confirmatory, corrective and improving examples. And what is true for a general term should presumably also be true for other expressions, as we already saw regarding proper names.

Indeed, if all we have in these cases is a shared convention, then a psychological element needs to be involved, even if only in an implicit way, constituting what could be called a *non-reflexive cognitive application of the rule*. Definition (C) makes explicit a convention normally instantiated in our heads as unreflected tacit application, whenever we make conscious use of the word 'chair,' which only confirms the traditional standard explanation.

10. Challenges to the traditional explanation (ii): Gareth Evans

There is another argument against the claim that we have tacit cognitive access to semantic conventions that govern our use of expressions. This argument comes from the philosopher Gareth Evans, who directly influenced McDowell. Evans invites us to contrast a person's belief that a substance is poisonous with a mouse's disposition not to consume it. In the case of a human being, it is a genuine belief involving propositional knowledge; in the case of a mouse, it is a simple instinctive disposition to react in a certain way to a certain smell, not a true belief. Proof of the difference is the fact that:

> It is of the essence of a belief state that it be at the service of many distinct projects, and that its influence on any project is mediated by other beliefs. (Evans 1985: 337).

If someone believes a certain substance is poisonous, he can do many different things based on that belief. He can test his belief by feeding it to a mouse, or if he is depressed, he can try to commit suicide by swallowing

a dose. He can also relate his belief that the substance is poisonous to a variety of other beliefs. For instance, he might believe he will become immune to a poison by consuming small amounts every day, gradually increasing the dose... As our knowledge of semantic rules is not susceptible to such inferences, thinks Evans, it consists not of actual belief states, but rather of *isolated states*, not very different from those of the mouse. Therefore, they are not cognitive (or pre-cognitive) psychological states in a proper sense of the word. (Evans 1985: 339)

The characterization of belief proposed by Evans is interesting and in my view correct, but his conclusion does not follow. Certainly, it agrees with many of our theories of consciousness, according to which a belief is only conscious if it isn't insular, while an unconscious belief is insular – though there are degrees of insularity. But the crucial point is that Evans' argument blinds us to the vast gulf between our semantic uses of language and the mouse's behavioral disposition to avoid consuming poison.

As a weak but already useful analogy, consider our knowledge of simple English grammar rules. A child can learn to apply these rules correctly without any awareness of doing so, and some adults who have never learned grammar are also able to apply these rules correctly to many different words in many different contexts. Moreover, even if our knowledge of these grammar rules is very often unconscious, with sufficiently careful examination we can often bring them to consciousness.

The point becomes still clearer when we consider our simple example of an implicit semantic-cognitive rule, the criterial rule (C) for the application of the concept-word 'chair' to the identification of chairs. Certainly, a person can derive many conclusions from this rule. He can predict that five persons cannot sit side-by-side on a single chair. He knows that one can transform a chair into a stool simply by cutting off its backrest. He can know the price and if he could buy a similar chair. He knows that by standing on a chair, he can reach an overhead ceiling lamp… He knows all this and much more, even without ever having consciously considered definition (C). And this only means that we can have a belief state enabling us to identify chairs, putting it at the service of many different projects mediated by other beliefs without being explicitly aware of the involved meaning-rule (C).

We can see a continuum, beginning with more primitive and instinctively determined dispositions and ending with semantic-cognitive rules of our language and their effects. It includes dispositions like those of mice, which cannot be cognitive, because they are instinctive (it is utterly implausible to think that a mouse could be reflexively conscious). There are also more sophisticated ones, like our unconscious beliefs, thoughts

and cognitions, which we can consciously scan and reflexively access (presumably through metacognitive processes).

If we accept the view that our semantic rules are usually conventional rules exemplified in the simplest cases by models like (C), then we must reject the radicalism of positions such as those of Evans and McDowell. After all, the application of such rules allows us to make many different inferences and relate them to many other conceptual rules. Rule (C) has greater proximity to the rules of English grammar than to the innate dispositional regularities demonstrated by a mouse that instinctively avoids foods with certain odors. Moreover, it is clear that in such cases, unlike the mouse, for people inferences to other beliefs are always available. This can be so even if we admit that our semantic-cognitive rules do not in themselves possess the widest availability proper to those completely conscious belief states considered by Evans.[18]

The root of the confusion is that the semantic rules in question, with and because of their apparent triviality, have not yet been investigated in a sufficiently systematic way. In an academic milieu dominated by science, the procedure that leads to their discovery does not seem worthy of careful investigation. Nevertheless, to proceed more systematically in this seemingly trivial direction is in fact philosophically invaluable, and this is what I will do in the remainder of this book.

11. Unreflected semantic cognitions

I believe contemporary theories of consciousness support the traditional view according to which we have implicit knowledge of our meaning-rules. I will begin by appealing to *reflexive* theories of consciousness. But first, what are these theories?

In the philosophical tradition, the idea of reflexive consciousness was already suggested by John Locke with his theory of *internal sense* (1690, book II, Ch. 1, §19). Reflexive theories of consciousness were introduced to the contemporary discussion by D. M. Armstrong (1981: 55-67; 1999: 111 f.). We can summarize Armstrong's view as saying there are at least two central meanings of the word 'consciousness.' The first is what he calls

[18] Freud distinguished (i) unconscious representation able to associate itself with others in processes of unconscious thought from (ii) unconscious representation that remains truly isolated, not associated with other representations, which for him would only occur in psychotic states and whose repression mechanism he called *exclusion* (*Verwerfung*). Evans treats the relative insularity of our non-reflexive awareness of semantic rules in a way that suggests Freud's concept of exclusion.

perceptual consciousness, which consists *in the organism being awake, perceiving objects around it and its own body*. This is the simplest sense of consciousness. John Searle wrote that consciousness consists in those subjective states of sentience or awareness that begin when one wakes up in the morning after deep, dreamless sleep and continue throughout the day until one falls asleep at night, or lapses into a coma, or dies (2002: 7). By this, he meant chiefly perceptual consciousness. This is also a very wide and consequently not so distinctive sense of consciousness since less developed species also have it. For instance, we can say that a hamster sedated with ether loses consciousness because it ceases to perceive itself and the world around it. We are surely justified to assume that when a hamster is awake it has some primitive form of cognition of the world around it, as shown by its behavior. However, this excessive extension of perceptual consciousness only contributes to its irrelevance. We are aware of the world in the same way a hamster seems to be conscious of it but in a much more sophisticated, more human sense of the word. Certainly, a mouse perceives a cat, but it is unlikely to know it is facing its archenemy. This also holds for inner feelings. A snake may be able to feel anger, but we hardly believe a snake is aware of this anger since it certainly has no reflexive consciousness.

Now, what distinguishes a mouse's perceptual awareness and a snake's anger from our own conscious awareness of things around us and from our own feelings of anger? The answer is given by a second sense of the word 'consciousness' which Armstrong considers the truly important one. This is what he termed *introspective consciousness* and that I prefer (following Locke) to call *reflexive consciousness:* This is a form of consciousness that we can define as the *reflexive awareness of our own mental states.*

According to one of Armstrong's most interesting hypotheses, reflexive consciousness emerges from the evolutionary need of more complex systems to gain *control* of their own mental processes by means of higher-order mental processing. In other words: our first-order mental events, like sensations, feelings, desires, thoughts, and even our perceptual consciousness of the world around us, can become objects of simultaneous introspections with similar content (D. M. Rosenthal called these metacognitions *higher-order thoughts*[19]).

[19] *Cf.* Rosenthal 2005. In this summary, I will ignore the dispute between theories of higher-order perception (Armstrong, Lycan) and higher-order thought (Rosenthal), and still others. In my view, David Rosenthal is right in noting that Armstrong's perceptual 'introspectionist' model suggests the treatment of cognitions of a higher-order as if they contained *qualia* and that it is implausible that higher-order processes have phenomenal qualities. Armstrong, on his side,

According to this view, only when we achieve reflexive consciousness of a perceptual state can we say that this state 'becomes conscious' in the strong sense of the word. So, when we say in ordinary speech that a sensation, a perception, a sentiment or a thought that we have 'is conscious,' what we mean is that we have some kind of *metacognition* of it. This shows that Armstrong's perceptual consciousness is actually a kind of *unconscious awareness*, while reflexive consciousness – the true form of consciousness – is probably a faculty possessed only by humans and a few higher primates such as orangutans.[20]

Now, let us apply this view to our tacit knowledge of semantic-cognitive rules. It is easy to guess that we usually apply these rules without having a metacognitive consciousness of them and therefore without making ourselves able to consciously scrutinize their structure. In other words, we apply these rules to their objects cognitively[21], and these rules deserved to be called 'cognitive' because they generate awareness of the objects of their application. But in themselves these rules usually remain unknown, belonging to what I above called unconscious awareness. Hence, it seems that we need to resort to some kind of metacognitive scrutiny of our semantic-cognitive rules in order to gain conscious awareness of their content.

One objection to using this kind of theory to elucidate tacit knowledge of our rules is that there are a number of interesting first-order theories of consciousness that do not appeal to the requirement of higher-order cognition. In my view, we can classify most, if not all, of these apparently competitive for theories as *integrationist* theories of consciousness. We

---

seems to be right in assigning a causal controlling role to higher-order introspection, since for him consciousness arises from the evolutionary necessity to maintain unified control over more complex mental systems. Aside from that, although Armstrong doesn't use the word 'thought,' he would certainly agree that there is some kind of higher-order *cognitive* element in the introspection of first-order mental states, an element that interests us here. I prefer the term *metacognition* for these higher-order cognitions since I believe that not only Rosenthal but also Armstrong would agree that introspection is a cognitive phenomenon.

[20] I will pass over the traditional idea that of themselves first-order mental states automatically generate metacognitions. This view would make it impossible to have perceptual consciousness without introspective consciousness. However, this view not only seems to lack a convincing intuitive basis; it also makes the existence of unconscious thoughts incomprehensible.

[21] Some use the term 'pre-cognitive' for what is implicitly known. I use the word 'cognitive' in a broader sense, including the cognition of what is implicitly, unconsciously known.

can do this because they share the idea that consciousness of a mental state depends on its degree of integration with other mental states constituting the system. This is certainly the case of Daniel Dennett's theory, according to which consciousness is 'brain celebrity': the propagation of ephemerally fixed contents influencing the whole system (1993, Ch. 5). This is also the case with Ned Block's view, according to which consciousness is the availability of a mental state for use in reasoning and directing action (1995: 227-47). This is likewise the case with Bernard Baars' theory of consciousness as the transmission of content in the spotlight of attention to the global workspace of the mind (1997). And it is also the obvious case of Giulio Tononi's theory, according to which consciousness arises from the brain's incredible capacity to integrate information (2004: 5-42). These are only some well-known contemporary first-order theories of consciousness that are historically consonant with Kant's view, since according to him in order to be consciously recognized, a mental state must be able to be unified (integrated) into a single Self. From the perspective of such integrationist theories, an unconscious mental state would be one that remains to a greater or lesser extent *dissociated* from other mental states. And all these views seem to some extent plausible.

The objection, therefore, would be that I am trying to explain tacit knowledge of language by relying solely on metacognitive theories of consciousness, ignoring all others. However, I believe there is more than one way around this objection. My preferred way is the following: we have no good reason to think integrationist and reflexive views of consciousness are incompatible. After all, it makes sense to think that a mental state's property of being the object of metacognition also seems to be a condition – perhaps even a necessary one – for the first-order mental state to be more widely available and more easily integrated with other elements constituting the system. As Robert Van Gulick wrote in the conclusion of his encyclopedia's article on consciousness:

> There is unlikely to be any theoretical perspective that suffices for explaining all the features of consciousness that we wish to understand. Thus, a synthetic and pluralistic approach may provide the best road to future progress. (2014)

Indeed, we can reinforce our suspicion by reconsidering a well-known metaphor developed by Baars: A conscious state of mind is like an actor on stage who becomes visible and therefore influential for the whole system, because he is illuminated by the spotlight of attention. However, I think it seems reasonable to think that this could happen only because of the operation of some sort of searchlight of the will added to some sort of

metacognitive mental state constituting the true spotlight. Hence, one could easily argue that the first-order mental state is accessible to the rest of the system and hence conscious *due* to its privileged selection by some kind of metacognitive state of attention.

My conclusion is that our awareness of semantic-cognitive rules and the possibility of scrutinizing them metacognitively would be able to resist integrationist theories, since they also leave some room for conscious processes able to be scrutinized by means of reflexive attention. Consequently, assuming some kind of reflexive *plus* integrationist view, the plausible conclusion remains that we can have cognitive states that make us conscious of their objects even if such states are not in themselves proper objects of consciousness. Thus, it seems plausible that only if the first order cognitive processes are objects of the (reflexive, metacognitive) scrutiny of attention can we subject them to conscious analysis. And most of our semantic-cognitive rules belong to such cases.

It seems to me that this kind of assumption could explain why we can have non-conscious or implicit or tacit cognitions when we consciously follow semantic-cognitive rules without being cognitively aware of the content of these rules and consequently without being able to analyze them. They remain implicit because we do not pay attention to these rules when we apply them and because even when this occurs, they are not there as objects of cognitive scrutiny. These rules are there, to use an old metaphor, like spectacles. When seeing things through them, we are normally unaware of the lenses and their frame. Assuming this kind of view, we conclude that we can distinguish two forms of cognition:

(i) *Non-reflexive cognition*: This is the case with cognitions that are not conscious, because they are not accessed by a higher-order cognitive process and/or focused on by inner attention, etc. (e.g., my perceptual consciousness when I use rule (C) for identifying a chair.)

(ii) *Reflexive cognition*: This is the case of cognitions accessed by a higher-order cognitive process and/or focused on by inner attention, etc., being for this reason able to be the object of conscious access followed by reflexive scrutiny. Any mental states, sensations, emotions, perceptions, and thoughts can be called reflexive if they are accompanied by higher-order cognitions and/or focused on by inner attention. (This is a previous condition needed for the kind of reflexive scrutiny that can make us aware of the semantic-cognitive rule (C) for the identification of a chair as requiring a seat with a backrest, designed for use by only one person at a time.)

Once in possession of this distinction, we can better understand the implicit or tacit status of the cognitive senses or meaning-rules present in uses we make of expressions. When we say that the structures of semantic-cognitive rules determining the references of our expressions are often implicit (as in the case of the semantic rules defining the words 'chair' or 'Aristotle'), we are not assuming that they are properly pre-cognitive or definitely non-cognitive, lacking any mentality. Nor that they are completely isolated or dissociated from any other mental states (in the latter case, we would lack even the ability to choose when to apply them). What we mean is just that the cognitive instantiations of these conventional rules are of a non-reflexive type. That is, although consciously used (we know we are using them), they are not likely to be the subject of some form of reflexive cognitive attention. Moreover, as already noted, there is a reason for this, since the structures of these rules are not the focus of our attention when we use the corresponding concept-word in an utterance. By uttering our sentences our real concern is much more practical, consisting primarily in the cognitive effects of applying these rules. As an obvious example: if I say, 'Please, bring me a chair,' I don't need to explain this by saying, 'Please, bring me a non-fixed seat with a backrest, made to be used by only one person at a time.' This would be discursively obnoxious and pragmatically counterproductive: it would be almost impossible to communicate efficiently if we had to spell out (or even think of) all such details each time we applied semantic-cognitive rules. What interests us is not the tool, but its application – in this case, to inform my hearer that I would like him to bring me a chair. In linguistic praxis, the meaning isn't there to be scrutinized, but instead to be put to work.

A consequence of this view is that in principle our inner attention must be able to focus on non-reflexive semantic-cognitive rules involved in normal uses of words and scrutinize them metacognitively by considering examples of their application or lack of application. Taking into consideration the variable functions and complexity of our semantic-cognitive rules enables the philosopher to decompose them analytically into more or less precise characterizations. It seems it is by this mechanism, mainly helped by examples, counterexamples, comparisons, and reasoning, that we become aware of the conceptual structure of our philosophically relevant expressions.

12. Conclusion

Summarizing this chapter, we have found two main devices for methodological orientation: (A) the primacy of *established knowledge* and (B) the method of *philosophizing by examples*. They will be used as guides in this book. Particularly relevant in this context is the idea that we can still see philosophy as an analytical search for non-reductive surveillable representations of our natural language's central meaning-rules. It is almost surprising to verify that more than two-thousand years after Plato we still have reasons to accept the view that solving some of our most intriguing philosophical problems would require only deeper and better analyzed explanations of what some central common words truly mean.

APPENDIX TO CHAPTER II

MODAL ILLUSIONS: AGAINST SUPRA-EPISTEMIC METAPHYSICAL IDENTITIES

> *Die Probleme, die durch ein Mißdeuten unserer Sprachformen entstehen, haben den Charakter der Tiefe. Es sind tiefe Beunruhigungen; sie wurzeln so tief in uns wie die Formen unserer Sprache, und ihre Bedeutung ist so groß wie die Wichtigkeit unserer Sprache.*
> [The problems arising through a misinterpretation of our forms of language have the character of depth. They are deep disquietudes; they are rooted as deeply in us as the forms of our language, and their significance is as great as the importance of our language.]
> —*Wittgenstein*

> Philosophy unties the knots in our thinking, which we have tangled up in an absurd way; but to do that, it must make movements that are just as complicated as the knots.
> —*Wittgenstein*

Although exceedingly original and thought-provoking, Saul Kripke's philosophical application of modal logic to problems of reference is in my view burdened by a disturbing web of confusion. Since many would disagree, I will give a short justification of my conclusion by critically discussing his article 'Identity and Necessity' (1971), which preceded the more developed views defended in his book *Naming and Necessity* (1980), since it takes his central ideas, as it were, directly from the oven. The paragraphs below summarizing Kripke's article are in italics, in order to distinguish them from paragraphs containing my own comments. After my comments on this article, I provide an Addendum containing a series of brief criticisms of positions taken by Kripke, Hilary Putnam, Gareth Evans, David Kaplan, Tyler Burge and John Perry, as part of my project of debunking their metaphysics of reference and meaning.

Kripke begins by considering the modal argument for the necessity of identity statements. This argument can be summarized as follows. Given the principle of indiscernibility of the identical, according to which (x) (y) ((x = y) → (Fx → Fy)), and given the principle of identity, according to which (x) □(x = x), we can conclude that if the property F must necessarily be applied to x, then y must also have this property. That is, it is necessary *that y equals x. In symbolic notation, (x) (y) (x = y) → (□(x = x) → □(x = y)), namely: (x) (y) (x = y) → □(x = y).*

*This apparently inconsequential formal result leads Kripke to the bold conclusion that identities between proper names are necessary. We know this by universal instantiation □(x = y) → □ (*a = b*). That is, if* a *and* b *are real names and a = b is a true identity, then this identity is necessarily true. This would concern identities like 'Hesperus is (the same as) Phosphorus' and 'Cicero is (the same as) Tulli': they must necessarily be identical. Further, if F and G are* theoretical predicates*, defined as essential designators of properties, if they form a true theoretical identity of the form (x) (Fx = Gx), then this identity is also necessarily true. That is why identities like 'Heat is (the same as) molecular motion' and 'A state of mind is (the same as) a physical state,' if true, are necessary.*

Kripke recognizes that identities between names and between theoretical identities have generally been considered contingent. There seem to be good reasons for this. Consider the statement 'Hesperus is Phosphorus.' Since Hesperus is Venus seen at dusk (evening star), and Phosphorus is Venus seen at dawn (morning star), it was an important astronomical discovery that they are actually the same planet, as Frege pointed out. Therefore, this seems *not to be a necessary, but rather a* contingent *empirical truth. The same applies to theoretical identities such as 'Heat is molecular motion.' This identity resulted from scientific discovery and could be false because if caloric theory (the theory that heat consists of a self-repellent fluid called caloric) were correct, heat would not be molecular motion. This seems to be a contingent statement since it clearly could be otherwise.*

Kripke's thesis, however, is that contrary to appearances, all these identities, despite having been discovered a posteriori, are necessary, even if they do not seem *to be: they are* necessary a posteriori identities. *To reinforce his thesis he introduces a famous distinction between the* rigid designator*, here defined as* a term that refers to the same object in any possible world where this object exists or would exist,[1] *and the* non-rigid

[1] Later he generalized, writing that a rigid designator 'in every possible world designates the same object,' which includes worlds where the object does not exist

or accidental *designator, which* can refer to different objects in distinct possible worlds (1971: 146). *Proper names and terms of natural species are rigid designators referring to the same object in different possible worlds. Most definite descriptions, by contrast, are* accidental designators, *designating different objects in different possible worlds. An example of an accidental designator would be the definite description 'the inventor of bifocals,' which in our world refers to Benjamin Franklin, but in some possible worlds could refer to any other person or even to no person. In contrast, the proper name 'Benjamin Franklin' always refers to the same person in any possible world where Benjamin Franklin exists. Thus, if we have an identity in which the identity symbol is flanked by proper names, this identity is* necessarily true *if true at all, considering that these proper names, being rigid, must have the same bearers in all possible worlds where their bearers exist.*

It is clear that a mathematical term can be seen as a rigid designator, insofar as its application does not depend on how the world is. But is it really impossible for proper names to be other than rigid designators? In an attempt to show that Kripke is wrong and that sometimes they could be accidental designators, we can imagine the following. Suppose it were discovered that shortly after G. W. Bush's childhood an extra-terrestrial being took possession of his body, assumed his identity and impersonated him from then on, subsequently being elected president of the United States and performing all the actions attributed to him. In this case, wouldn't the proper name 'G. W. Bush' be unwittingly used to refer to this extra-terrestrial being instead of to the son of Barbara and George Bush, who was born on 6.7.1946, becoming in this way an accidental designator?

The idea that a proper name is a rigid designator could easily withstand objections like that. According to Kripke, the reference of a proper name is due to an act of *baptism* followed by a causal-historical chain. But this only means that the true G. W. Bush, as the bearer of the rigid designator 'G. W. Bush,' would long since have ceased to exist. On the other hand, the embodied extra-terrestrial being, whose true name was, say, Gkw9, would have had its proper first baptism in some remote place and time. Hence, the name G. W. Bush (in fact here a mere alias of Gkw9) would apply to this same extra-terrestrial being in any possible world where he existed, serving here as a homonymous rigid designator. With the same

---

(1070: 48). However, this last view directly contradicts the meaning of the verb 'to refer', which can only be rightly applied when the referred to object exists (*Cf.* Searle 1969: 77; Plantinga 1974: 80).

symbolic form (G. W. Bush), we have in fact two contextually distinguishable proper names that are two different rigid designators with different bearers and nothing more.

It is important to remark that if we applied the neodescriptivist theory of proper names summarized in the Appendix of Chapter I, the results would be the same. Accordingly, the proper name's bearer is the object that satisfies its identification rule. What this identification rule requires is that this object sufficiently and better than any other satisfies the inclusive disjunction of the fundamental description-rules, which are the localizing and the characterizing rules. Regarding the adult G. W. Bush (as Gkw9), for instance, the localizing description includes his earlier spatiotemporal career on another planet before his embodiment on Earth, and after that his service as President in Washington and his subsequent life. On the other hand, the characterizing description would include his main accomplishments, including his election as 43rd president of the USA, leading the country after 9/11, beginning wars in Iraq and Afghanistan, together with his earlier accomplishments on a distant planet as Gkw9... In every possible world where the identification rule is adequately satisfied, G. W. Bush (as Gkw9) would exist. Hence, the identification rule for the name is also a rigid designator. Something of the kind could also be easily established for the child really baptized G. W. Bush, born on the Earth on 6.7.1946 who had that tragic fate… also making this name a rigid designator of another bearer by satisfying its own identification rule.

In addition, Kripke expects us to believe he has warranted the necessity of the identity between proper names by having discovered some radical metaphysical difference between proper names, on the one hand, and definite descriptions, on the other. What his words suggest is that a proper name could be attached to its reference without intermediaries by means of a direct (in my view mystical) relation instituted by the act of baptism. For him, this act does not really depend on any properties of the object, even if we are helped by their descriptions to identify it. Notwithstanding, he believes that this baptism allows the post-baptismal production of external causal-historical chains between speakers and hearers. These chains ultimately enable any speaker uttering the name as the last link of a chain to *refer* to the name's bearer.[2] A definite description, in contrast, is only an accidental designator. It would refer to different objects in different possible

[2] Rejecting the view of a particular as a bundle of abstract properties, Kripke concludes: 'What I do deny is that a particular is nothing but a "bundle of qualities," whatever that may mean' (1980: 52). He was certainly unaware of the then only recently introduced trope theory. (See Appendix of Chapter III)

worlds, presumably because it would have a completely different reference mechanism, based on what John Stuart Mill called a 'connotation,' defined by him as 'the description's implication of an attribute that the object may have' (1881, I, Ch. 2).

As I already noted, Kripke's explanation for this modal dichotomy, suggesting a difference in the nature of each referring process, is as illusory as it is dispensable. The only way to explain the modal dichotomy seems to be by appealing to the already exposed meta-descriptivist theory of proper names, which gives a really plausible justification for the contrast between the rigidity of proper names and the accidental character of their associated definite descriptions (See Appendix of Chapter I, sec. 7, 8). As we saw, the application of meta-descriptivist theory shows that descriptions are rigid only insofar as we compare them to the reference of the proper names they are associated with, which means that definite descriptions lacking an associated proper name are rigid. After such explanations, the idea of a rigid designator, at first seemingly so profound, turns out to be nothing but a technical formulation of an in the end much more trivial idea. It is the idea that *if we assume the existence of a proper's name bearer (that is, if we consider any possible world where this bearer exists) this proper name* must *be able to refer to it.* But no one would disagree with this!

Furthermore, unlike Kripke's view, the necessity of the rigid designator is here the product of *de dicto* conventions. I say this in agreement with John Searle's brilliant demystifying analysis of the *de dicto/de re* distinction (1983: 208-220). According to him, so-called *de re* beliefs are only a *sub-class* of *de dicto* beliefs, so that there can be no irreducible *de re* beliefs, as Kripke supposed. As Searle notes, beliefs are *de re* only in the sense that they are intended to refer to real objects, not that they harpoon real objects. Although there is a class of beliefs whose explanation depends on contextual characteristics, one should not equivocally conclude that such characteristics cannot be *entirely* represented as part of the intentional (mental) content! Under this assumption, the true difference between beliefs called *de dicto* and ones called *de re*, turns out to be a mere difference between *reports*. In a *de dicto* belief like 'Ralf believes that the man with the brown hat is a spy,' we commit ourselves only to the report of Ralf's belief. In a *de re* belief like 'About the man with the brown hat, Ralf believes he is a spy,' we also commit ourselves to the existence of the man with a brown hat. Hence, there is no reason why both beliefs at bottom should not be *de dicto* beliefs. Now, if we reject irreducible *de re* beliefs, we feel ourselves free

to reject the supra-epistemic metaphysical *de re* necessities assumed by Kripke.

As it was noted in the Appendix of Chapter I (sec. 6), my proposed metadescriptivist view of proper names makes them rigid designators because in any possible world where the proper name has a bearer, at least one combination of fundamental descriptions that allows its reference in accordance with its identification rule must be sufficiently and majoritarily satisfied. However, the reason for this rigidity is not metaphysical. It is simply because the identification rule *is what defines what any bearer of the proper name can be*. Now, considering identity between different proper names in statements of the form $a = b$ and assumed the proposed metadescriptivist theory of proper names, we may have two clearly different cases. The first is the following:

(A) Two different proper names of the same object have two different identification rules that identify their bearer under different guises, under different modes of presentation, simply because they take into consideration different perspectives in which different descriptions or groups of descriptions are satisfied. In this case, even being in themselves rigid designators, identifying the same object in all possible worlds as their own objects, we cannot without additional information conclude that they in fact refer to *the same* object in any possible world. Here it is an empirical matter to decide. We still do not know whether the identification rules of two names could be part of a common, wider identification rule and a new rigid designator having these rules as guises since we still do not have established this last rule. Consequently, in a first moment the empirical finding of an identity statement of the kind a = b would be seen as *contingent a posteriori.* The modal form of this identity would be $\Diamond$ (a = b). This was the case before astronomy showed beyond reasonable doubt that the morning star is the evening star, for instance, when for the first time someone observed the evening star in the sky all through the night and realized that it seems to be always the same as the morning star. (Venus cannot be tracked each night; it disappears for earthly observers during part of the year, when it passes behind the Sun, what must have given place for doubts).

The second case is the following:

(B) Often, after many and varied empirical experiences we have no reasonable doubt that a = b. In this situation, we establish a *new convention* – a rule according to which the different modes of presentation a and b, *the two different identification rules, are blended into a single identification rule with two different guises,* each of them emphasizing a different aspect or mode of presentation of the same object. In this case, however, what we ultimately have is *a single rigid designator* able to identify the same object in any possible world, even if under at least two different guises. The identity resulting from the newly established convention is then considered *necessary a priori.* Its modal form is □ (a = b) or □ (a[b] = b[a]). This is the case today, when we identify the morning star with the evening star, having as a conditional background our modern knowledge of astronomy. It is important to notice that at no point in this process do we need to resort to a Kripkean *necessary a posteriori identity*, unless we confuse the a posteriority of case (A) with the necessity of case (B).

Summarizing: The two initially independent identification rules are made constituents of only one more complex identification rule that includes both anterior rules. One could write this identity as 'Morning star [-evening star] = evening star [-morning star],' differing only by different emphases on each side of the identity sign. Anyway, the names 'morning star [-evening star]' and 'evening star [-morning star]' are used here as rigid designators for the same object, the planet Venus. Moreover, they are conventionally assumed to be *de dicto* rigid designators. Whether they are also metaphysically *de re* rigid designators, above and beyond any convention, is something that, it seems, no human being has the power to know.

Kripke also considers the problem of apriority. A priori truths are ones we can know without appealing to experience. Many consider the necessary *and the* a priori *to be equivalent. However,* for him*, the concept of* necessity *is* metaphysical – *about how the world* must be – *while the concept of* a priori *is* epistemic – *about how we* know *the world. Kripke thinks the two classes are not equivalent. Consider, he writes, Goldbach's conjecture that any natural number is the sum of two primes. This may be a necessary truth without the possibility of our knowing it a priori. In this case, it would have metaphysical necessity.*

The claim that necessity is metaphysical while apriority is epistemological seems to me not fully mistaken, but requires better specification. I reject this distinction in Kripke's formulation. His understanding would be justified only if we could discover real metaphysical *de re* necessities, since a *de dicto* necessity follows from a more trivial, conventionally established apriority, even if rooted in experience. Moreover, it seems to me that the awareness of the existence of metaphysical *de re* necessities in Kripke's sense is something that goes far beyond what our cognitive faculties could reach, simply because our empirical knowledge is inherently *fallible* – a point that has been consistently emphasized by philosophers of science from C. S. Peirce (1991, Ch. 7) to Karl Popper (1989, Ch. 10). From this perspective, the most we can do is *to postulate as natural laws those empirical regularities that are not only strongly inductively grounded, but also* the most deeply entrenched ones, *in the sense that they are strongly inferentially integrated with our most plausible system of scientific and modest commonsense beliefs.*[3] We cannot speak of a natural law's necessity going beyond this well-grounded postulation, since to prove this metaphysical necessity we would need absolute knowledge – something our epistemic fallibility precludes. Therefore, the so-called necessity of natural laws and what follows from them is simply a result of a well-grounded decision to *treat* them as necessary, and since this conventional decision is well grounded by deep entrenchment, we have a right to expect (*pace* Armstrong[4]) that they will resist counter-factual situations. These assumed necessities are necessities in a secondary sense of the word, of course, since they can possibly be denied without contradiction. However, once we postulate these necessities, we have a right to treat them as what we have made of them: rules of our own conceptual system. This explains why we constantly use derived statements of necessity like 'It is *necessary* to have fire to light a candle.' Such natural necessities should be epistemically identified with practical certainties, once we see that they can be treated as certainties,

[3] This high level of entrenchment seems to me the relevant reason we distinguish between regularities that have the 'necessity' of natural laws and those that are merely coincidental. This entrenchment creates the impression that our knowledge of natural law is of something that exists by logical necessity. (For similar approaches, see Tugendhat & Wolf 1983: 253; See also Mackie 1974.)

[4] D. M. Armstrong defended the view that scientific laws are necessary because they are relations between universals, which explains their resistance to counter-factual examples (2010, Ch. 5). However, the ontological price of this view seems simply too high.

insofar as we can grant them a sufficiently high degree of probability to put them beyond reasonable doubt.

Finally, we must ask what remains of the empirical root, the seemingly unknowable real objective essences responsible for 'metaphysical necessity.' One possibility is that it still has a function in a sense that recalls what Kant called an *idea of reason*. It seems that we have a directive idea (whose supposed reference is cognitively impossible to find), constructed only to offer a horizon able to measure and pragmatically motivate our investigation. This would be only a directive idea of a metaphysically *de re* necessity, which could justify our search for generalizations, as approximations of a mere *ideal* of an absolute, unquestionably necessary empirical knowledge. This ideal seems to serve as an in fact unreachable target, able to make possible the comparison between our approximations, allowing us to establish comparative degrees of assurance between our judgments. In this context, ideas like that of a *real essence* serve as heuristic tools, even if they cannot be true objects of reference. We proceed *as if* something achieved were objectively necessary. And in a similar way, we also proceed as if we knew ultimate truth and as if we had ultimate knowledge, at least until we discover that we have made a mistake.

Summarizing the profession of faith of the apparently old fashioned empiricist that I am: I admit that necessity might be a metaphysically loaded concept. However, it works for us as a conventional *de dicto* necessity which we can only believe to be rooted in some *de re* necessity, in a way similar to the way we can only believe that a nominal essence is rooted in some real essence. Anyway, it is not a *de re* necessity that can be epistemologically spelled out in the forms of a priori knowledge usually expressed.

If this approach to necessity is accepted, one could go ahead in suggesting a very broad distinction between two main kinds of necessity, both of them conventional and with essentially epistemic (and only ideally metaphysical) import:

(A) *Unconditioned necessities.* These are the formal necessities that we find in logic and mathematics, along with linguistic conventions. Their statements are clearly analytic and their negations are, regarding the system of signs to which they belong, contradictory or inconsistent.

(B) *Conditioned necessities*. These are necessities arrived at a posteriori, depending on empirical experience to be achieved. Because of this, they are not necessities in the strict sense of the

> word intended by Kripke when he speaks of metaphysical necessities because after being inductively or hypothetically-deductively reached, they are simply *conventionally postulated and assumed as necessities*. This is a very common sense of the term, though a weaker one, insofar as it is circumstantial, presupposing something like the truth of a theory or system or cluster of empirical beliefs sustaining it, e.g. the nomological necessity expressed in a statement such as 'Necessarily $V = \Delta D/\Delta t$ [assuming traditional kinematics].'

Although diversely defined, these two general kinds of necessity have a long tradition in philosophy that began with Aristotle. For him (A) was *absolute necessity*, the necessity in the fundamental sense, understood as the cause of itself, as in the case of the principle of non-contradiction (1984, vol. 2, *Met.* 1015b10). On the other hand, (B) would be a *hypothetical necessity*, a derived form, a necessity due to an external cause, like the necessity of water for sustaining life (1984, vol. 1, *Phys*. 2, 9, 200-230). The first was a necessity in the proper sense, since the opposite of it would imply a contradiction. The second, so-called hypothetical necessity, is such that its opposite does not imply a contradiction, or implies a contradiction *only under a given condition*, like the assumption of a belief-system or a linguistic praxis. Both A and B are conventionalized in the innocuous sense that they depend on conventions with varying degrees of arbitrariness.[5] Although I do not intend to elaborate this point here, my impression is that Kripke tends to oversee the distinction between A and B, treating weakened conditioned necessities as if they were unconditioned.

As for Goldbach's conjecture, the fact that it may be a necessary truth without our being aware of it does not mean that in this case its suggestion that any natural number is the sum of two primes is not an a priori truth, since it can also be an a priori truth without our being aware of it, insofar as we see an a priori truth as a truth that *can* be known without experience. It might also be necessary but unknown, insofar as it can be also a priori

[5] I think that Wittgenstein would classify as (B) conditioned necessities those implicit conventions that he called 'grammatical rules,' grounding useful linguistic practices (1984a). Here is his suggestion, in which I read the word 'rule' as a conventional (a priori) conditional necessary proposition: 'Every empirical proposition can serve as a *rule* if it is fixed as the immovable part of a mechanism, in such a way that the entire representation revolves around it, making it part of a system of coordinates independent of the facts.' (1984e, part VII: 437)

but unknown, presently being for us only possibly necessary and only possibly a priori. If it happens that we never discover its truth a priori, we will also never discover its necessity. And it is not impossible that someone will find a proof of this conjecture, finally giving it its cognitive status of a theorem with *a priori necessity*. Indeed, it is because mathematicians maintain the heuristic rule that it is possible to reach such an a priori necessity that they still insist on searching for proofs.

The most striking and revealing example of a necessary a posteriori *statement introduced by Kripke is that of the wooden lectern in front of him. He starts with the question: could this lectern have consisted, since the beginning of its existence, of ice from the Thames? Certainly not: It would be a different object. Thus, the statement 'This lectern, if it exists, cannot be made of ice,' is a necessary truth known a posteriori. Lecterns are usually not made of ice. This lectern seems to be made of wood, and it is not cold. Hence, it is probably not made of ice. Of course, this could be an illusion. It could actually be made of ice. But that's not the point, writes Kripke. The point is that* given the fact *that the lectern is not made of ice, but of wood, one cannot imagine that it could be made of ice.* Given the fact *that it is not made of ice, he concludes, it is* necessary *that it is not made of ice. More precisely: being P = 'This lectern is* not *made of ice,' and considering that we know both the a priori truth that 'If P then □P' and, from empirical research, that P is true... Kripke constructs the following argument, applying* a modus ponens*:*

(A)
1 $P \rightarrow \Box P$
2 $\underline{P}$
3 $\Box P$

It is, therefore, necessary that the lectern is not *made of ice, although this is only known a posteriori, through empirical research. The statement* 'This lectern is not made of ice' *is* a striking example of a *necessary a posteriori*!

Unfortunately, there is a well-hidden mistake in Kripke's argument. It concerns the epistemological status of P ('This lectern is not made of ice') in the second premise. In this premise, the truth of P is affirmed in complete disregard for the fact (earlier confusingly introduced by him) that P, like any empirical statement, can only be known to be true by inevitably *fallible* epistemic subjects. However, if this is so, then P can in principle

be false. In order to show my point clearly, I first need to define a statement as *practically certain* if it is *so likely to be true that we can ignore the probability of its being false*. This is usually the case when we can assign a statement a probability of being true very close to 1.[6] On the other hand, I define a statement as *absolutely certain, if it logically cannot be false*, having *a probability 1 of being true*, which makes it obviously necessary.[7] Considering this, we can instead say that statement P of the second premise should more precisely be written as (2'): 'It is *practically certain* (that is, it has a probability very near 1 of being true) that P.' Indeed, (2') must be true, because we know this. However, only God – the infallible and omniscient epistemic subject – could know with *absolute certainty* the truth of statement P (that is, God would be able to assign it the probability 1), warranting the factual existence of P. He would in this way give the state of affairs described by P a truly metaphysically *de re* necessity. Unfortunately, we cannot appeal to God in this matter… All we can know is that P is *practically certain* in the already stated sense. If we assume that all available information is true, then it is sufficiently likely that we can accept it as true. This must be so if we accept the fallibility of our empirical knowledge, its lack of absoluteness. (Not impossible is a radically skeptical scenario in which Kripke believes he knows that he is standing before a hard wooden lectern, which is supported by all available testimony and possible empirical tests, and, nevertheless, the lectern is really made of ice.[8])

Assuming this, consider Kripke's premises again. First, it is fully acceptable that *given the fact that* P, P follows by necessity.[9] So, what P → □P says is, '*If* it is *really* the case that P, then it is necessary that P,' and this, I concede, is a logical truth. However, what the antecedent of P → □P requires is that P implies □P only if P is really the case, which demands that our knowledge of P must be *absolutely* certain, with a probability of 1 and not just an assertion that a fallible knower 'holds to be true.' Only

---

[6] The concept is relevant for those who accept that (empirical) knowledge is *justified true belief*, because then they will need practical certainty regarding the conditions of truth and, consequently, knowledge itself. (*Cf.* Costa 2014, Ch. 5).

[7] I assume that 'P is necessary' means the same as 'P has probability 1.' Seen as a probability, the idea of a necessity without any epistemic import appears to be nothing but an empty fetishism of necessity.

[8] For a discussion of skeptical hypotheses, see Ch. VI, sec. 30.

[9] This works well and trivially with *formal* necessities. If P were 5 > 3, one could argue (1) 'If it is the case that 5 > 3, then it is necessary that 5 > 3. (2) It is the case that 5 > 3. (3) Therefore, it is necessary that 5 > 3.

when P has a sheer probability 1 of being true is it an unconditioned necessary truth.

In other words, only an absolutely certain truth would warrant the necessity of the consequent, which would require an infallible being as its knower. Hence, the most complete analysis of premise (1) must be (1'): 'If it is absolutely certain that P is the case (if P has the probability 1), then it is necessary that P.' Surely, premise (1) could not be analyzed as (1'') 'If it is practically certain that P is the case (that is, if P has a probability close to 1), then P is necessary,' because the mere probability of P, no matter how high, as it is less than 1, would not warrant the necessity of P. Once we admit the change of premises (1) to (1') and (2) to (2'), Kripke's argument can be made completely explicit as saying:

(B)

1' If it is absolutely certain (with probability 1) that P, then it is necessary that P.

2' It is practically certain (with a probability close to 1) that P.

3' It is necessary that P.

Obviously, argument (B) is *not valid*, since the *modus ponens* cannot be applied to (1') and (2') to give us (3'). The reason is that the antecedent of (1') does not mean precisely the same thing as (2'), which makes the argument equivocal, hence fallacious. I conclude that under more careful scrutiny Kripke's argument is clearly flawed and consequently insufficient to convince us that the utterance 'This lectern is not made of ice' is a metaphysically *necessary a posteriori* truth.

Now we can easily see a reason for Kripke's misleading claim that the conclusion of his argument must be *necessary a posteriori*. He ignores the fine semantic differences made explicit in version (B) of his argument, and in doing so he jumps to a conclusion that unduly joins the necessity of his argument's first premise with the aposteriority of its second premise, interpreting the flawed conclusion (3) as a *necessary a posteriori* truth, in which 'necessity' is used in the unconditioned sense.

Kripke then goes on to the analysis of identities between proper names such as 'Hesperus is Phosphorus' and 'Cicero is Tulli.' These empirical identities were traditionally seen as contingent. However, for Kripke they are identities between rigid designators, which makes them necessary identities, since in the most diverse possible worlds these names will refer to the same object, which would not be possible if Hesperus weren't Phosphorus or if Cicero weren't Tulli. We could, he says, have identified

Hesperus and Phosphorus with two different celestial bodies, but in this case the sentence 'Hesperus is Phosphorus' would have a different meaning…

In order to demonstrate that the statement 'Hesperus is (the same as) Phosphorus' cannot be necessary a posteriori, here we can produce an argument analogous to the argument applied by Kripke to the indexical predicative case of the wooden lectern. Calling Hesperus *h* and Phosphorus *p,* we can construct the following Kripkean *modus ponens*:

(h = p) → □ (h = p)
h = p
□ (h = p)

The Kripkean conclusion of this argument is that 'Hesperus is Phosphorus' would be an (unconditioned) necessary identity that has been reached a posteriori.

Nonetheless, here as well the *modus ponens* does not apply because although the first premise is true, the second premise would only conjoin with the first one to reach the conclusion '□ (*h* = *p*)' *if* it were able to give us an absolute certainty that 'h = p.' However, empirically this cannot be the case. In order to obtain absolute certainty (probability 1) that 'h = p' is the case, which enables us to reach the conclusion of the conditional, this truth must be discovered, not by inevitably fallible human epistemic subjects only capable of practical certainty, but again only by God, the omniscient and infallible epistemic subject.[10] Because of this, 'h = p' can be seen here as merely an empirically reached fallible conclusion, stating that it is practically certain (sufficiently probable) that 'h = p,' which is still far from absolute certainty or probability 1. The following reformulation demonstrates the argument's hidden flaw:

If it is absolutely certain (with probability 1) that *h* = *p*,
then □ (*h* = *p*).
It is practically certain (with a probability close to 1) that *h* = *p.*
□ (*h* = *p*)

[10] God would be the only being able to know created things in their metaphysical necessities *de re*, perhaps because he knows them by sustaining them in their existence.

Since we cannot have the absolute certainty required by the identity of the antecedent of the first premise with the second premise, the equivocal character of the argument becomes clear. We cannot use the *modus ponens* to derive the a posteriori necessity of $h = p$. In this interpretation, the statement 'Hesperus is Phosphorus' is *contingent a posteriori*. It cannot be metaphysically necessary, because since this identity is only highly probable, it will always be possible that Hesperus is not Phosphorus. For instance, although extremely unlikely, it isn't logically impossible that the gods have until now maintained an incredibly complex illusion of knowledge in human minds, and that the planets are nothing more than a swarm of fireflies that assemble every night to decorate the celestial Vault. In this case, when seen by the naked eye, Hesperus would have a different location than Phosphorus, but it would appear identical to Phosphorus when viewed through a telescope – not because it is the same planet or even a planet at all, but as an effect of witchcraft. (The identity would be falsified, but together with it as well our whole astronomical system of beliefs. This once more shows that we can defend a conditional natural necessity regarding an identity statement, since by assuming its grounding system its opposite is made inconsistent.)

Kripke's second example is very different, and one should not confuse it with the first one. It concerns the utterance 'Cicero is Tulli.' Assuming my proposed neodescriptivist theory of proper names, the localizing description for his identification is (concisely) 'the person born in Greece on March 1, 106 BC and deceased in Rome on July 12, 43 BC,' while the characterizing description is (concisely) 'the most famous Roman orator, also a statesman, jurist, and philosopher.' His whole name was 'Marcus Tullius Cicero.' Since the proper name is not a fundamental description, but rather an auxiliary one (he could easily be given another name in a different possible world), Kripke is only relying on the fact that not all speakers know that Cicero and Tullius are parts of the same proper name, as a convention in our own world. The statement informs the hearer that the same bearer of the fundamental descriptions implied by each term flanking the identity sign is referred to by only part of the same person's whole name.

The result is that the statement's aim turns out to be a trivial one, namely, to communicate to the hearer a convention regarding the auxiliary description 'the person whose name was "Marcus Tullius Cicero".' Hence, the right answer is that 'Cicero is Tullius' only communicates part of a *necessary a priori linguistic convention*, once the convention that the whole name is 'Marcus Tullius Cicero' is decided a priori, just as is the convention that a triangle is a trilateral figure. Moreover, to say that the

statement 'Cicero is Tullius' is a posteriori would be to confuse its belonging to a definition in our actual world – which is a question of being informed about conventions – with the possible names that the same reference could have been given in different counter-factual situations. Indeed, it is possible that Cicero could have been given the name 'Marcus Titus Cicero' in a different possible world and even in ours, making the identity 'Cicero is Tullius' false. However, this is as trivial as to say that in a very different language (or world) people use a different name for 'triangle,' for instance, 'colmio.'[11] Consider the statement found in a bilingual dictionary, 'triangle means colmio.' It is not, say, an a posteriori inductive result of experience. It is the obvious expression of a *necessary a priori* identity regarding conventions.

The next of Kripke's examples concerns the identity between kinds *of things, as in the already discussed statement* '*Heat is molecular motion.*' *Many think that this identity, being the result of empirical research, expresses an a posteriori truth. However, for Kripke this is* a necessary a posteriori identity *because the heat (in a gas)* cannot *be anything other than molecular kinetic energy, since the terms 'heat' and 'molecular motion' are rigid designators. It may be, he says, that the Earth could at some time have been inhabited by beings who feel cold where we feel hot and vice versa, so that for them heat would not be identical with molecular motion. However, this would not be the case, since we understand heat as the sensations caused* in us *by this molecular motion.*

The fact that the two nominal terms flanking the identity sign are rigid designators does not warrant that they are rigid designators of the *same* bearer, picking up the same object in the same possible worlds, since any identity can be false (e.g., '8 = 8' is true, '8 = 9' is false, though these numbers are all rigid designators). Thus, the claim of identity alone warrants nothing.

On the other hand, as noted in the Appendix to Chapter I, since we have ways to translate rigidity in descriptive terms for proper names, we have reasons to guess that the same can be done with general terms. That is, we could link the two ascription rules for heat in gas and kinetic molecular energy to create a unified ascription rule that has two different guises – two different but interchangeable main designative criteria, producing a *necessary a priori* identity.

[11] 'Colmio' means triangle in Finnish.

On the other hand, assuming the independence of the identification rules, we can employ the same strategy used above in order to discredit Kripke's view that 'Heat is molecular motion' is a case of a *necessary a posteriori*. Thus, considering heat in gas and kinetic molecular energy as rigid designators that necessarily designate one essence, we could construct the following Kripkean argument calling heat in gas H and kinetic molecular energy M:

$(x)\ ((Hx = Mx) \rightarrow \Box\ (Hx = Mx))$
$\underline{(x)\ (Hx = Mx)}$
$(x)\ \Box\ (Hx = Mx)$

Clearly, the same problem reappears. The first premise says only that *given that* (or *if*) the identity $(x)\ (Hx = Mx)$ is really the case, then it is *necessarily* the case that all heat is molecular motion. Or, in the epistemic parlance, if it is *absolutely certain* that all heat in gas is kinetic molecular energy, then it is *necessary* that all heat in gas is kinetic molecular energy. However, since the identity affirmed in the second premise, being empirical, is inevitably fallible, the following paraphrase of the above argument is inescapable:

(x) If it is absolutely certain (with probability 1) that $(Hx = Mx)$,
then $\Box(Hx = Mx)$.
(x) It is practically certain (with a probability close to 1) that $(Hx = Mx)$.
(x) $\Box(Hx = Mx)$

Here again, the more explicit formulation shows an equivocal and consequently fallacious argument for the same reason given in the above cases. It is thus clear that we cannot in this way conclude that the statement 'Heat (in gas) is the same as molecular motion' is a Kripkean *necessary a posteriori* truth. Thought of in this way, it is a *contingent a posteriori* truth.

The last of Kripke's examples should be the most important one. It is intended as a refutation of identity theories of the mind-body relation, according to which 'Pain is (the same as) such and such a brain state' would be a contingent a posteriori scientific discovery that has not been made. But, as Kripke writes, 'pain' and 'such and such a brain state' are rigid designators here, for they refer to essential properties. However, if this is the case, the identity theorist is in trouble, because this identity should be necessary*, which frontally clashes with the fact that* whenever

you feel pain you do have pain, *while no one is denying that it is possible to conceive that we feel pain without having the corresponding brain states. For a theistic philosopher like Kripke this makes identity theory implausible.*

I find this argument puzzling. First, as a matter of fact, one can feel pain without there being an identifiable physical cause, for instance, in the case of hypnotized subjects who feel imaginary pains, and in many other cases. However, even if we ignore this, admitting that we cannot consciously feel legitimate physical pain without having some qualitative subjective state of pain, the fact that we can *conceive* of pain without a corresponding brain state does not prove anything. Similarly, the fact that Descartes could imagine his mind existing without his body could not prove that a mind can exist without a body.[12] Why does this force us to think that a future neuroscience might not be able to show us that by speaking of such and such a brain state we make a rigid reference to exactly the same thing we experience as a state of pain, so that this identity would then be established as *conditionally necessary*, in a way similar to the case of heat as molecular kinetic energy?

It is true that feeling pain isn't the same as detecting heat outside us by feeling hot inside. Only the feeling of being hot inside is subjective and immediate like pain. But in the same way that a Martian might feel cold when we feel hot, a Martian might feel a tingling sensation when we feel pain. And we can similarly imagine that feelings of pain, like those of heat, could in some way be interpersonally identified in the brain using suitable technical procedures (*Cf.* Ch. III, sec. 13). Hence, the only real difference that remains between the two cases is that kinetic molecular energy in gas is located externally, outside of a person's head, while such and such a brain state is located internally, within a person's head. But why should it be relevant to show that pain isn't the same as neuronal behavior, and even *necessarily* the same in a conditioned, non-metaphysical sense of the word?

Kripke concludes his argument by saying: '...heat is picked out by the contingent property of being felt in a certain way; pain, on the other hand, is picked out by an essential property' (1971, note 18). However, even in the case of pain there is no certainty that the feeling of pain, if it is put into words, picks out the real essential property of pain, just as there is no guarantee that a discovered general neuronal pattern of pain picks out the

[12] In the last case, this is because 'imagine', like 'doubt' and 'conceive,' are verbs of propositional attitudes, which do not allow extensional inferences.

real essential property of pain, beyond a new kind of nominal essence. The identity can be stated as real only from the hypothetical perspective of a conditionally established necessity. To see this, imagine a world where most people's pain is imaginary and it is extremely easy to mistake imaginary for real physical pain; worse than this, try to imagine a tribe a people whose pain is always imaginary, but so well justified that we all mistakenly believe their pain to be real.

As I see it, in most cases Kripke confuses the *a posteriori* element of a *contingent a posteriori* discovered identity with its well-grounded conventional establishment as something *necessary a priori* – a *de dicto* necessary truth. This leads him to believe in a *supra-epistemic de re metaphysical necessity* which is discovered a posteriori. In doing so, he assigns to ontologically unknowable identities the same status of epistemologically assumed identities. He proceeds as if we could assert ontological (metaphysical) truths without considering our epistemic capabilities and their intrinsic fallibility. He refuses to accept that we can never completely separate the epistemic from the ontic; and in so doing, he denies an insight that marks the foundation of modern philosophy by Descartes, namely, that we lack access to supra-epistemic truths.

Addendum: disposing of externalism

A great variety of arguments developed by Kripke and other externalist philosophers merit closer examination. In what follows, I will limit myself to a few comments, since a more detailed analysis would far exceed the scope of this book.

1. There are a variety of supposed examples of *necessary a posteriori* truths that were later proposed by Kripke and others. I will consider as example (i) 'Cats are animals' (Kripke 1980: 181-2) because unlike the above examples, it is no identity statement. For Kripke this is a necessary statement since we cannot conceive of a cat that is not an animal; but it is also a posteriori, since it was discovered by means of empirical experience. Therefore, it is a *necessary a posteriori* truth.

Now, for me, this is all wrong. Statement (i) can be interpreted in two ways, depending on the context. Here they are:

(a) *Contingent a posteriori*: a primitive tribe that sees a cat for the first time might easily suppose, based on its aspect and behavior, that it is an animal like others. The tribe arrives at this knowledge a posteriori, because it is based on experience, and contingently, because it is liable

to revision since, for them, it could be that the cat is, in fact, a legendary forest spirit, that merely chooses to assume the form of an animal.

(b) *Necessary a priori*: Zoologists, accepting the truth of our contemporary taxonomy, according to which the cat is classified as an organism belonging to the *Animalia* kingdom, assume that the statement (i) is *necessary a priori*. So interpreted, the statement is necessary in the conditioned or hypothetical sense of the word, and it is a priori (analytic) because it abbreviates tautology (ii) 'Animals called cats are animals.'

One can only arrive at the necessary a posteriori by confusing the conditioned necessity of interpretation (b) with the a posteriori character of interpretation (a).

2. Another form of *necessary a posteriori* later suggested by Kripke concerns origins. For him, rigidity makes true parenthood necessary. He considers the case of Queen Elizabeth II (1980: 112 f.). Indeed, she would not be Queen were she not the daughter of Albert, Duke of York, and his wife, Lady Elizabeth Bowes-Lyon.

This is a suggestive, but biased example, since in the case of a queen the ovum origin acquires maximal importance, which we would easily analyze as a case of contamination of the identification rule by descriptions of origins (*Cf.* Appendix to Ch. I, sec. 9 (v)). Suggestiveness and biased concrete examples work here as a way to confuse things and mimic a false sort of relevance. In the case, Elizabeth became Queen of England because her uncle abdicated the throne, making Elizabeth's biological father the new King, thereby establishing her as the biological heir to the throne. A similar case is the necessity of the origin of the species homo-sapiens as a result of evolution from such and such previous species of hominids (McGinn 1976) – an empirical discovery that can achieve definitional status. But from another perspective, precisely identical exemplars of homo-sapiens fabricated in a future biological laboratory could be devoid of any proper necessity of origin, due to the indirectness and irrelevance of a possible link between such clones and our ancestors. Anyway, there is no reason to see the association of natural necessities with proper names as more than a well-grounded *de dicto* necessity established by us.

By contrast, consider the statement (i) 'Ishmael Lowenstein is the son of Abel and Berta Lowenstein.' According to a Kripkean philosopher, this statement should be *necessary a posteriori,* because even if it is known a

posteriori, an adult with different parents stemming from a different ovum and a different sperm cell would not be Ishmael Lowenstein.

However, suppose that the adult Ishmael makes the shocking discovery that his parents are not his biological parents. He was mistaken for a different infant in the hospital where he was born, and a subsequent DNA analysis showed that he was actually the son of Amanda and Mario Belinzoni, whose supposed son was baptized with the name Carlos. Of course, this is no reason to think that Ishmael thereby ceases to be Ishmael or should be renamed Carlos. This name is even printed on his birth certificate and driver's license. If asked, he could insist on answering that his name is Ishmael Lowenstein, certainly with the agreement of others who know him. This is consistent with our identification rule, since Ishmael still satisfies the localizing and characterizing conditions sufficiently better than does any other person.

Nevertheless, our conclusion might also be less straightforward. Consider again the complete statement (i) 'Ishmael is the son of Abel and Berta Lowenstein,' addressing the question of parenthood.[13] One could use as a criterion of parenthood those who cared for the child and raised him with loving care until adulthood. In this case, we can regard the statement 'Ishmael is the son of Abel and Berta Lowenstein' as true, even though he was conceived from Mario's sperm cell and Amanda's ovum. So understood, the statement 'Ishmael is the son of Abel and Berta Lowenstein' could be better seen as *contingent a posteriori*. Contingent because it could be false that they cared for and nurtured him; a posteriori because knowledge of this kind is acquired through experience.

Notwithstanding, it is not difficult to imagine a situation in which Kripke's view would apply. Suppose we were in Nazi Germany, the Lowensteins were Jewish, and the Nazis had arrested the family. For the Nazis the criterion of parenthood was clearly biological. In this case, if the Nazis were well informed about the mix-up of babies, Ishmael Lowenstein would be considered the son of Mario and Amanda Belinzoni, while Carlos would be considered the true son of Abel and Berta Lowenstein, and as such should be arrested and sent to a concentration camp… With regard to the proper name, however, the matter isn't so simple. Nonetheless, it is conceivable that the Nazis decided to establish a rule according to which a person's true and legal name must be the name linked to his biological origin so that they would replace the name of

[13] Today there are at least five competing theories of parenthood (genetic, labor-based, intentional, causal and pluralistic ones), and there is no consensus on the right cluster of criteria (*Cf.* Brake & Millum 2016, sec. 4).

Ishmael Lowenstein with Mario Belinzoni, and the unfortunate young man called Mario Belinzoni would become Ishmael Lowenstein. In this case, the statement (ii) 'Carlos Belinzoni (Ishmael Lowenstein) is the son of Mario and Amanda Belinzoni' isn't a *necessary a posteriori* truth. It could be seen (a) as a very probable contingent a posteriori discovery, insofar as one emphasizes the *fact* that the name Carlos Belinzoni (= Ishmael) should now mean the same thing as the son of Mario and Amanda as an (a posteriori) discovered truth and a (contingent) conclusion reached inductively. On the other hand, it could be seen (b) as emphasizing the stipulated decision to treat 'Carlos Belinzoni' (= Ishmael), as an abbreviation of 'the son of Mario and Amanda Belinzoni…' making this an essential part of his identification rule, so that statement (ii) would be seen as a conditioned *necessary a priori.*

I guess that Kripke would answer these objections by noticing that what we discover about parents doesn't matter. What matters is that *if one were born to parents x and y, then one could not have been born to any other parents* (1980: 113). But so understood, this is a trivial *a priori* statement like 'If a woman is wearing a hat, she must have something on her head.' The problem is that any attempt to give a concrete example will lose its supposed character of metaphysical necessity, since it will be based upon unavoidably limited empirical experience, being therefore in principle fallible. Indeed, there is only one way in which a given origin would generate a *necessary a posteriori*, namely when viewed by an infallible knower. God would discern that Ishmael is definitely the son of Mario and Amanda Belinzoni, giving this the probability 1 of necessity independently of experience, and he would know it as a *de re* metaphysical necessity. He would know this as the only being able to achieve supra-epistemic knowledge. We, as fallible knowers, do not possess this gift. By using concrete examples, Kripke gives the impression of having made a metaphysical discovery about the world when he is really only saying something that is either trivial or contains an anticipation of what can be easily derived from the already proposed neodescriptivist view of proper names.

3. Worse than the *necessary a posteriori* is Kripke's later invention, the *contingent a priori* (Kripke 1980: 54-56). It uses a case involving the platinum rod stored in Paris, once designated as the standard metric unit of length. According to Kripke, analysis of meaning is something different from definition; the first is necessary, the second is not (although he gives no satisfactory justification for this). Then he claims that the definition of 'one meter' as '*the* length of S at t_0' is not *necessary a priori*, but

contingent a priori! The reason is that the term 'one meter' is a rigid designator, while 'the length of S at t_0,' being a definite description, is an accidental designator, only helping to fix the reference. The accidental designator can change its reference. For example, in different possible worlds the length of S at t_0 could be greater or less than a meter on Earth, for reasons such as heating or cooling. Thus, in another possible world (in a contra-factual situation) one meter could be a length different from 'the length of S at t_0.' Consequently, the statement 'the Paris platinum rod is one meter long (has the length S),' although established a priori, is contingent.

This argument could be adequate if we accept the existence of some metaphysical reason for the distinction between names as rigid designators and descriptions as accidental designators. However, the real reasons for the distinction are non-metaphysical, as I think I made clear enough in the Appendix of Chapter I (sec. 8-9): definite descriptions are only accidental when dependent on proper names, but not when they make proper names depend on them, as in the present case. Consequently, we have reason to doubt Kripke's affirmation that after being established definitions are neither meaning-giving nor necessary. For it seems clear that the definition of a meter as 'the length of S during Δt'[14] is a *stipulative definition* made to establish the proper meaning of one meter. Thus, why cannot 'one meter' have been chosen as a mere abbreviation of 'the length of S during Δt,' *whatever this length is?* Why cannot 'the length of S during Δt' be a rigid designator, no less than is 'average kinetic molecular energy'? Assuming this, our intuitive reasoning would be to think that whether the length of the standard meter changes or not, in its function as a standard of measurement, the meter remains the same, since the standard meter is *defined* as being necessarily whatever length S has in the Δt when it is used as a standard. This means that in any possible world where the standard meter exists, the length of this meter will continue to be considered the same, no matter what its cross-world comparative length may be.

It is only for practical reasons that, wishing to preserve the comparative function of measuring length, it is better for us to use the most rigid and most unchangeable possible standard meter. Suppose, by contrast, that the standard meter were a kind of very elastic rubber rod, continually changing its length. It would remain the same standard meter, of course,

[14] The symbol 'Δt' is more correct. The rod served as a standard not only at t0, but rather during the entire period in which it was conventionally designated to have its function.

but it would be quite impractical as a model for measurement. Using this elastic standard in accordance with the given definition, we could be forced to accept the absurd conclusion that a woman who two hours ago was 1.67 m tall is now 2.24 m tall; or that objects with very different sizes could be the same size if we measured at different times…

The point is that if you accept that the statement 'A meter = the length of S during Δt, whatever length it has when measured' is the actual definition of a standard meter – and it really is – this definition given by the definite description 'the length of…' isn't contingent, but *necessary*, since it is a convention that cannot be falsified in any possible world where it holds. Moreover, this definition is *a priori*, for we do not need to have any experience in order to know its truth. Consequently, the following identity can be considered the right definition of a meter:

> One meter (*Df.*) = the length of the standard rod S during any moment of Δt, disregarding the possible world-circumstance in which its length is permitted to be effectively considered.

The *definiendum* 'one meter' is nothing but an abbreviation of the *definiens*. Like any stipulative definition, this identity is *necessary and a priori*. The description that constitutes the *definiens* is rigid because we have established it as applicable in any possible world where the standard rod S exists, and the name 'one meter' is rigid because it abbreviates the characterization rule established by the *definiens*.

4. Another attempt to exemplify the *contingent a priori* could come from Gareth Evans' example using the name 'Julius,' which he arbitrarily stipulates as naming 'the inventor of the zipper' (1982: 31). According to some authors, the statement (i) 'Julius was the inventor of the zipper,' is *contingent a priori*. It is a priori because we do not need experience to know this; but it is also contingent, since it is possible that 'Julius' e.g., sustained brain injuries when very young and grew up too retarded to invent the zipper (Papineau 2012: 61).

Concerning statement (i), we again find a dual reading:

(a) On the one hand, it is *contingent a posteriori*. It is contingent because in a counter-factual situation it could be that the zipper was not invented by anyone or that it was invented by several persons… but it is also a posteriori, because its truth depends on experience to be discovered.

(b) On the other hand, we could paraphrase 'Julius invented the zipper' as (ii) '*Assuming* that the zipper was invented and that only one person invented the zipper, we have decided to call this person "Julius".' However, this paraphrase of (i) is not contingent a priori, but *necessary a priori*. It is a conditioned necessary harmless stipulation that is a priori because it is established independently of experience.

We conclude that neither in case (a) nor in case (b) do we have a *contingent a priori*.

5. A curious attempt in the same direction was given by the following utterance: 'I am here now,' proposed by David Kaplan (1989: 509). This would also be a *contingent a priori truth*. It is said to be a priori because each of its terms directly refers respectively to the *agent*, the *place* and the *time* of a given context of utterance. This excludes the possibility of its falsity. However, since we can imagine counter-factual circumstances in which I would not be here, this utterance is only contingently true.

This example is also deceptive. Even 'I am here now' can be a false statement in our real world. I remember a case related by Dr. Oliver Sacks of a patient who had a seriously deranged perception of temporal continuity. Because of this, her daily life was a succession of time-lapses: she could think 'I am here now,' imagining that she was still in her bedroom, when in fact she was already in her kitchen. Thus, in this case, 'I am here now' is intuitively felt by us as *empirically false*! This shows that the statement 'I am here now' can be in fact *contingent a posteriori*, since it is falsifiable and since we need to know the real context of the experience to determine its truth.

6. I also disagree with Hilary Putnam's view, according to which the meaning of the word 'water' must fundamentally be outside our heads.[15]

This is perhaps the most influential argument for semantic externalism. According to Putnam's Twin-Earth fantasy, in 1750 Oscar1 on the Earth and his *Doppelgänger* Oscar2 on Twin-Earth – two almost identical planets with the same history – both simultaneously saw water and called what they saw 'water.' Since the chemical structure of water wasn't yet known in 1750, all that Oscar1 and Oscar2 could have had in their heads

[15] I say 'fundamentally,' because Putnam admits that surface descriptions (stereotypes) and classifications (semantic markers) are internal secondary mental features of meaning (1975: 269).

was the idea of a *watery fluid* (a substance that at room temperature is liquid, transparent, odorless, tasteless…). However, in the hypothetical case, they were actually referring to very different compounds, Oscar1 to H_2O, while Oscar2 was referring to XYZ. Water on Twin-Earth (in the example) has a very different chemical composition, summarized by Putnam as XYZ, even though it has the same appearance and properties of water on our Earth. For Putnam this proves that the *meaning* of water – a word that for him essentially refers to quantities of molecules with the same microstructure of H_2O – could not have been in the Oscars' heads, since in their heads they had *the same state of mind*, namely, the idea of a watery fluid and nothing more. Putnam's slangy comment is the most famous externalist statement: 'Meaning just ain't in the head.'(1975: 227) As he summarizes in a central passage:

> Oscar1 and Oscar2 *understood* the term 'water' differently in 1750, although they had the same psychological state, and although, given the state of development of Science in their epoch, the scientific community would need to take circa 50 years to discover that they *understood* the term 'water' differently. Hence, the *extension* of the term 'water' (and, in fact, its *meaning* in the pre-analytic intuitive use of the term isn't a function of the speaker's psychological state). (1975: 224; my italics.)

That our understanding and meanings are not in our heads is a shocking conclusion, later radicalized by John McDowell's inference that even the *mind* must be external to the head because it is the locus of our manipulation of meanings (1992: 36).

My neodescriptivist answer is that Putnam's result is due to his overlooking the fact that 'water' has two descriptive *nuclei* of meaning: a *popular* and a *scientific* one.[16] First, there is an old popular nucleus of meaning of the word 'water.' This nucleus is phenomenal and also dispositional and can be summarized by the expression 'watery fluid.' It is a substance that at normal temperatures is a transparent, odorless liquid that quenches thirst, can be used to wash things, is a universal solvent, extinguishes many kinds of fire, falls from the sky as rain, forms rivers, lakes and oceans, freezes when cooled below 0 degrees C, evaporates when heated above 100 degrees C, has high surface tension, etc. This was the well-known meaning until the end of the eighteenth century. Then a

[16] For a more detailed argument, including a more careful neodescriptivist analysis of the meanings of 'water', see Costa 2014, Ch. 3.

great semantic upheaval occurred. A new dimension of meaning was increasingly added: the scientific nucleus, which can be summarized as 'quantities of H_2O.' Water was discovered to be a chemical substance that results from combining hydrogen and oxygen, as summarized in the formula $2H_2 + O_2 = 2H_2O$, which can be shown by burning hydrogen mixed with oxygen and by electrolysis. Moreover, inter-molecular hydrogen bonds are responsible for water's high surface tension, liquid state at room temperature, etc. Both nuclei of meaning are intrinsically *inferential*. Nonetheless, they are also obviously objects of *descriptions* (since in opposition to anti-descriptivist bias, the domain of what can be described is much wider than a merely perceptual domain, containing descriptions of dispositions, micro-structures, etc.), which can be confirmed by consulting any good dictionary.[17] We use the word 'water' on an everyday basis in accordance with what we know from the inferential semantic rules of these two nuclei. Meaning is not a question of all and nothing.[18] Furthermore, it is easy to see that depending on contextual factors, one of these two clusters of meaning tends to come to the fore.

This summary already allows the following plausible internalist explanation of the Twin-Earth fantasy. First, in 1750 the two Oscars had in their heads only the nucleus of meaning expressed by 'watery fluid,' so that the extension and meaning of the word water were the same for both Oscars. However, when Putnam considers what is happening, he is overvaluing and unconsciously projecting the *scientific nucleus of meaning* of the word 'water' into the two Oscars' utterances, as if it were the only semantically relevant one. What he does then is simply to treat the two Oscars *as mere indexical devices for the projection of the new*

[17] For instance, the main definition in a *Merriam Webster dictionary* contains elements of *both* popular and scientific nuclei of meaning. It is the following: water = the liquid that descends from the clouds as rain, forms streams, lakes, and seas, and is a major constituent of all living matter and that when pure is an odorless, tasteless, very slightly compressible liquid oxide of hydrogen, H_2O, which appears bluish in thick layers, freezes at 0°C and boils at 100°C, has a maximum density at 4°C and a high specific heat, is weakly ionized to hydrogen and hydroxyl ions, and is a poor conductor of electricity and a good solvent. (On the descriptive relevance of the dispositional and scientific properties of water and their presentation in dictionaries, see Avrum Stroll, 1996: 71).

[18] We need to know only the most common descriptions, and this is enough for our adequate use of the word in more or less vague contexts. We do not need to know all the descriptions of water; even chemists do not know all of them. Did you know, for instance, that when water is cooled to near absolute zero (-273.15° C.), it changes again from a solid to a liquid state?

scientific nucleus of meaning, whose true locus is in fact our own heads/minds (i.e., those of Putnam and his readers), based on our knowledge that Oscar1 is pointing to H_2O, while Oscar2 is pointing to XYZ. Consequently, the different scientific meanings of the word 'water' are not in the world and outside of our heads, as Putnam believes. They are in Putnam's head when he thinks his thought-experiment, and in our heads when we read his texts. Today we all have some general knowledge about the scientific nucleus of meaning (summarized as H_2O) and may guess that a different scientific nucleus with similar effects (XYZ) would perhaps not be completely impossible. Finally, since Putnam and his readers have *different* scientific meaning-descriptions in their heads (H_2O and XYZ) when unconsciously projecting them (respectively) onto Oscar1 and Oscar2 by using them as indexical devices, these different meanings obviously remain, as they should, internal properties of minds. This also explains why (again helped by our instrumental referential devices called 'Oscars') we give them different extensions.

The neodescriptivist view suggested above leads us to see that the meaning of 'water' receives *variations of emphasis* according to with what we could call the *context of interest* in which a word is used, that is, the context of its circumstantial utility. In this case, there is a popular and a scientific context of interest leading to different interpretations as follows:

(a) In a popular context of interest (e.g., of fishermen who use water for cooking, drinking, and washing) the sense that is emphasized in the statement 'Water is H_2O' is that of a watery fluid. In this case, 'Water is H_2O' means above all (i) 'Watery fluid = fluid consisting of H_2O.' Taken at face value, this is a *contingent a posteriori statement.* Contingent because, at least in principle (though very improbably), it could be proved false; a posteriori because the conclusion is based on experience. Its modal form, though possessing a high level of probability, is still $\Diamond (a = b)$.

(b) In a *scientific context of interest* (e.g., in a chemist's laboratory) the scientific nucleus of meaning is emphasized. Here 'Water is H_2O' means above all (ii) 'Water as dihydrogen oxide = H_2O.' As expected, (ii) is a conditioned *necessary a priori* statement with the modal form $\Box (a = b)$, more conspicuously, $\Box (a[b] = b[a])$, assuming the truth of basic chemie. In this context, even if water were not a watery fluid, but rather something like a black oily fluid, it could still be called 'water,' insofar as it had the right microstructure.

Conclusion: the Kripkean classification of the statement 'Water is H_2O' as a *necessary a posteriori* statement results from a confusion between the a posteriori nature of statement (a) and the (conditioned) necessity of the similar statement (b). Since both senses are components of the whole meaning of 'water' and may alternatively come to the fore, it is easy to fall into a confusion resulting from lack of attention to the pragmatics of natural language, since Putnam and Kripke overvalued the scientific nucleus. We will deal with these kinds of confusion when we examine Wittgenstein's account of transgressing the internal limits of language. In this case, the confusion is a matter of *equivocity* resulting from the ill-fated attempt to import popular into scientific usage (*Cf.* Ch. III, sec. 11).

7. There are two other examples of Putnam trying to show that meaning is not only located in the external physical world, but also in *society*. In the first one, he assumes that aluminum and molybdenum are only distinguishable by metalworkers and that Twin-Earth is rich in molybdenum, used to manufacture pots and pans. In addition, he imagines that the inhabitants of Twin-Earth call molybdenum 'aluminum' and aluminum 'molybdenum.' In this case, he writes, the word 'aluminum' said by Oscar1 will have an extension different from that of the word 'aluminum' said by Oscar2, so that they mean different things with the word. However, as they are not metalworkers, they have the *same* psychological states. Hence, the meaning of these words is external to what happens in their heads, depending on their societies.

Our answer is the following. Consider how Oscar1 and Oscar2 use the words 'aluminum' and 'molybdenum.' They are not metalworkers, and what they have in their minds is indeed *the same thing.* It is as much so as the extension that they are able to give to their concepts of aluminum and molybdenum, which in the example includes both. For the metalworkers of Earth and Twin-Earth, on the other hand, aluminum on the Earth and the molybdenum of Twin-Earth (called by their inhabitants 'aluminum') have very different constituent properties, which means that metalworkers would have something very different in their heads. The Oscars may confuse both things, but only because they do not really know the intrinsic properties of these things and they are using the words in an incomplete, *subsidiary sense*. However, since *we* know the differences between the amounts of these metals on both planets, we can consider the aluminum and the molybdenum respectively observed by Oscar1 and Oscar2 and unconsciously take both persons *as referential devices for the different meanings we have in our heads.* In this case, we would say that Oscar2 is referring to what his linguistic community calls aluminum, but which in

our linguistic community is called molybdenum, while Oscar-1 is indeed referring to what we and our linguistic community call aluminum.

That people should use the words in accordance with the conventions of their linguistic community does not make the meaning external. It only makes it dependent on implicit or explicit agreements of members of their communities. In the two Oscars case, this agreement concerns only superficial properties. In the metalworkers' case, this agreement also concerns intrinsic properties. These agreements are always located in individual heads, even if differently distributed in heads belonging to a social network.

In his second example, Putnam considers differences between elm and beech trees. Most of us do *not* know how to distinguish between the two. However, we are able to guess correctly that these words are not synonymous, having different extensions, even without knowing the meanings of the two words. Hence, according to him the difference in meaning is not in our heads, but in society.

In response to Putnam, the important point to be noted is that most of us really do lack *sufficient* knowledge of the meanings of the words 'elm' and 'beech.' However, we already know something very generic about them: we surely know that they are trees, and we consider it probable (though not certain) that these two names refer to distinct kinds of trees.[19] With the help of these convergent descriptions (*Cf.* Appendix to Chapter I, sec. 5), we are able to insert these words into a sufficiently vague discourse. Moreover, we often do this while waiting for the suspected distinguishing information to be offered by specialists – those privileged speakers with *sufficient knowledge* of the meanings of these words. They are the only persons really able to identify examples of these different kinds of trees, so that without them these words would have no specific usage. The point is that meaning – sufficient or not – is always in the heads of speakers, even if (as I also agree) this meaning is located within

[19] In a later text (1988: 29), Putnam notes that if I know that a beech tree isn't an elm, I also know that an elm isn't a beech tree, which means that my knowledge is *symmetrical*, so that the representations are the same; furthermore, the words 'beech' and 'elm' are only phonetic shapes without meaning (1988: 27). But the semantic element here is just that we have reasons to believe that with these two names privileged speakers mean *different kinds of trees*, being able to detail the differences. Thus, that the description 'A beech tree is a tree that is different from an elm tree' is sufficient to allow us to insert these words in discourse as probably referring to different kinds of trees that after correction by others will be asymmetrically classified.

many heads that make up the communicative network of a socio-linguistic community.[20]

Concerning these two cases, Putnam appeals to a *division of linguistic labor* in order to account for the variety of meaning dimensions that may be possessed by different speakers. As he writes:

> We may summarize this discussion by pointing out that there are two sorts of tools in the world: there are tools like a hammer or a screwdriver, which can be used by one person; and there are tools like a steamship, which require the cooperative activity of a number of persons to use. Words have been thought of too much on the model of the first sort of tool. (1975, p. 229)

This is an important suggestion. However, it is far to confirm an externalist view of meaning. It is rather neutral. After all, the idea of a division of labor in language has already been suggested by internalist philosophers, from John Locke to C. S. Peirce (Smith 2005: 70-73). The former philosopher championed a theory of meaning as something consisting of internal psychological ideas. In effect, the division of labor is perfectly compatible with the fact that, even if socially shared, meaning remains in the heads of speakers, specialists or not, in different dimensions and degrees. In none of the above cases does meaning need to be located outside of heads.

Finally, to be fair, Putnam expresses himself much more cautiously in a later text (1988, Ch. 2), e.g., by suggesting that 'reference [as meaning] is *fixed* by the environment itself,' calling it 'the contribution of the environment' (1988: 32). However, we can read the word 'fixed' in two ways. In the first, we understand 'fixed' in the sense in which the external physical and social world is what ultimately produces referential meanings in our minds or heads. This is an obvious truism – something that a weak internalist (= a very weak externalist) like myself would be proud to agree. In the second way, which Putnam must intend to suggest, what he means with the word 'fixed' remains a too subtle metaphor to be intelligibly rescued, except by confessing that he is speaking about *reference* and not really about meaning. But one does not need to be a philosopher to know that references are in these cases obviously external, since belonging to the

[20] We can also find the right information in books, the internet, etc., but in order to be there, it must first in some way or measure be located in the human minds inside our heads...

external world. Putnam's externalism is an imaginatively brilliant philosophical effort that ends either in triviality or in confusion.

8. Now, I wish to reinforce my anti-externalist arguments discussing Tyler Burge's social externalism of thought, which in some ways complements Putnam's argument (Burge 1979). What Burge's text supports is the view that *the proper contents of thought or belief and propositional attitudes are external.*

I will first summarize Burge's argument and then show that it is easy to find a much more plausible weak internalist explanation of what happens, simply by elaborating a point already made by John Searle (2004: 284-6). In order to make it as clear as possible, instead of following Burge's counter-factual mental experiment, I will follow Searle's version. Suppose that a man named Oscar, residing in region A, feels pain in his thigh and therefore goes to see a certain Dr. Fugly, whom he tells:

(i) I think I have arthritis in my thigh.

Since arthritis is a painful inflammation of the joints, the doctor regards this belief as obviously *false*, since one cannot have arthritis in the thigh. Suppose that Oscar afterward travels to the very remote region B of his country and visits a certain Doctor Enoc because of the same health issue. But although in region A arthritis has its usual conventional meaning, in the remote region B people use the word 'arthritis' in a much broader sense, as referring to any kind of inflammation. Suppose that having forgotten his visit to the first doctor, Oscar once more tells this new doctor that he has arthritis in his thigh, having in mind exactly the same thing as previously. Now, in region B, as expected, the new doctor will confirm his suspicion, agreeing with Oscar's unquestionably *true* belief.

Based on a similar example, Burge's reasoning goes as follows. Without doubt, when Oscar claims he has arthritis in his thigh in both the first and second regions, his psychological states are exactly the same, just as his behaviors are the same. But the thought-contents expressed in the two utterances must be different, since thought-contents are truth-bearers, and the thought expressed in the first utterance is false, while the thought expressed in the second is true. However, the same thought cannot be both true and false! Moreover, in the second region the word 'arthritis' receives a new meaning, called by Burge 'tharthritis.' His conclusion is that the contents of the thoughts cannot be merely psychological. These contents must also belong to the outside world, to the social communities where the speakers live. (Burge 1976: 106)

Against this conclusion, it is not hard to find a commonsense internalist-descriptivist explanation for what happens. For a healthy weak internalism (that is, a minimalist externalism that admits that our mental subjectivity unavoidably depends on external inputs), in region B the concept-word 'arthritis' is the expression of an ascription rule constitutive of a meaning that is more general, designating any kind of inflammation. According to this rule, 'an inflammation that occurs in the thigh' serves as a criterial condition and belongs to the sense affixed to the word 'arthritis' in the linguistic community of region B. Thus, although the thought expressed in the sentence 'I think I have arthritis in my thigh' spoken by Oscar is precisely the same in the two linguistic communities, there is a fundamental difference that John Searle rightly identified as follows:

> Our use of language is *presumed* to conform to the other members of *our community*, otherwise we could not intend to communicate with them by using a common language. (2004, 184-5; my italics)

That is, when Oscar says to Doctor Fugly, 'I believe I have arthritis in my thigh,' he must assume that his ascription rule for the predicate 'arthritis' conventionally belongs to the language that other competent speakers of the language conventionally apply. The whole of what Oscar has in his mind (first actually and then dispositionally) in his utterance in the linguistic community of region A is:

(a) I have arthritis in my thigh… (and I am assuming that pain and inflammation in my thigh are accepted as a usual symptom of arthritis by the linguistic community of region A, to which my present interlocutor, Dr. Fugly, belongs).

This is *false* because the second sentence of the conjunction is false. Let's now see what is (first actually and then dispositionally) meant when Oscar tells the second doctor he has arthritis in his thigh:

(b) I have arthritis in my thigh… (and I am assuming that pain and inflammation in my thigh are accepted as a usual symptom of arthritis by the linguistic community of region B, to which my present interlocutor, Dr. Enoc, belongs).

Now statement (b) is *true*. Although the statement 'I have arthritis in my thigh' says the same thing, it has *a hidden indexical content that differs from (a) to (b).* However, this hidden indexical meaning *still belongs to*

Oscar's mind. Thus, it is true that if we confine ourselves to the content expressed by Oscar's thoughts when making the same utterance in both places, we see the statements as identical. However, there is an *overall* difference in what the hearers have in their minds (that is, in their heads) when hearing each utterance. It is different because Oscar wrongly assumed he was following conventions accepted by Doctor Fugly in the first linguistic community, while he later correctly assumes he is following conventions accepted by Doctor Enoc in the second linguistic community.

When he speaks with the doctor from community A, Oscar infringes on the principle that in order to achieve truth, verifiability rules constituting the content of thoughts should be in consonance with the conventions assumed by the linguistic community where the thoughts are expressed. But the correlative assumption isn't infringed on in community B, when Oscar speaks with Doctor Enoc. The conventional truthmakers given to members of the two social communities of speakers are different, although the semantic assumptions related to them by Oscar remain the same.

To be fair to Burge, we need to remember that he called attention to something important: the truth or falsehood of utterances depends on their conformity with linguistic conventions adopted by the speaker's community. This is already a relevant point, although it does not touch the claim that anything involved in thought-contents or beliefs (understood as senses or meanings of sentences) is outside the internal psychological realm, as it were in some mysterious way dispersed throughout the external socio-physical environment, as a strong externalist would like us to believe.

A final and more important point is the following: The given explanation allows us to make a healthy internalist paraphrase of the well-known distinction between *narrow content* and *wide content*. For the externalist point of view, narrow content is what is in the speaker's mind, while wide content is in some way external. A healthy internalist analysis of Burge's example allows us to propose that the *narrow* content of a thought restricts itself *to the semantic-cognitive verifiability rule that constitutes it.* This rule is expressed by the statement 'I think I have arthritis in my thigh.' On the other hand, the *wide* content of a thought is *what is indexically assumed in the speaker's mind as the adequate social convention that he expects to be satisfied by the narrow content.*

8. Finally, one word about John Perry's argument for the *essential indexical* (1979). I will be brief since I am repeating an argument I presented in more detail in another text (Costa 2014, Ch. 4). Contrary to Frege, Perry's view is that the senses of indexicals are inevitably linked

with the external circumstances of utterances, which can be proved by the fact that one cannot translate them into eternal sentences without any loss of meaning. The upshot is that, regarding indexicals, externalism of meaning is unavoidable.

In Perry's main example, he is pushing his shopping cart through a supermarket and notices that there is a trail of sugar on the floor. He begins to search for the source of the mess only to realize that he himself is the one who is spilling sugar on the floor. This leads him to say: (i) 'I am making a mess,' and as a result, he changes his behavior. Now, suppose we translate his statement into a non-indexical statement like (ii) 'Perry is making a mess.' This (almost) non-indexical statement cannot preserve exactly the same meaning. He could, for instance, be suffering from Alzheimer's disease, so that he has forgotten his name is Perry. In this case, he would know the truth of (i), but not the truth of (ii). The conclusion is externalist: no non-indexical statement is able to rescue the whole content of an indexical utterance. Some semantic content must unavoidably belong to the world.

However, I think there is, in fact, a way to preserve the whole content of the indexical, detaching it entirely from its indexical context. It is a technique I call *transplanting*: if you need to change the location of a plant, you almost never take the plant alone, but the plant together with the necessary amount of earth in which it is rooted… By analogy, here is how Perry's example appears after transplanting:

> (iii) At 10:23 a.m. on March 26, 1968, in the confectionery supplies section of Fleuty Supermarket in the city of Berkeley, CA, after noticing a sugar trail leading away from his shopping cart, Perry says *that* he is making a mess (or: 'I am making a mess').

What counts now is the truth of this eternal sentence[21] (iii) in which the indexical subordinate sentence is presented after a that-clause. Although containing indexical elements ('he' plus present tense), statement (iii) does not refer to the indexical context, since the indexical subordinate clause *refers indirectly*. It refers to what Frege called *the thought* (the belief-content) expressed by Perry in the subordinate clause that follows (a that-clause) or in the sentence with quotation marks (1892: 28). Thus, protected

[21] This might not be a perfect eternal sentence, but this does not change our conclusions since it is questionable if a statement without any kind of indirect indexical dimension is possible. If I say, 'The Earth is round,' I am already localizing the subject in our solar system. In this sense, all our empirical statements are indexicals.

by its surrounding description (the 'volume of earth' offered by the eternal sentence), the Fregean *sense* of 'I am making a mess' is here integrally transplanted without loss into the non-indexical context of a thought-content with a much wider reference.[22] What this argument shows is that the so-called essential indexical is not essential at all, since we can explicitly internalize its apparently external components.

[22] Phenomenal elements are obviously lost, but they do not belong to the conventionally grounded meaning. For a reconstruction of Frege's indirect reference in subordinate clauses, see the Appendix of Chapter IV, sec. 5 (iv).

CHAPTER III

WITTGENSTEINIAN SEMANTICS

> Philosophers constantly see the method of science before their eyes and are irresistibly tempted to ask and answer questions the way science does. This tendency is the real source of metaphysics, and leads the philosopher into complete darkness.
> —*Wittgenstein*
>
> *Im Anfang war die Tat.*
> [In the beginning was the deed.]
> —*Goethe*

My aim in this chapter is not so much to interpret Wittgenstein, as to reconstruct and sometimes develop his insights on meaning in a way that shows more coherence and relevance than we might suppose at first glance.[1] What I am seeking is something that in his own terminology could be called a *surveillable representation* (*übersichtliche Darstellung*) of the grammar of the concept-word 'meaning,' particularly concerning representative language. Before beginning, I would like to offer my views on something we could call the 'semantic-cognitive link.'[2]

---

[1] As will be clear, the assumption that guides my reconstruction is that Wittgenstein was not making repeated attempts to explain the nature of meaning that always ended in some kind of failure, erratically followed by new attempts, as some interpreters seem to believe. What he did was to develop different, often analogical approximations, each addressing the approximated issues from new perspectives, such suggestions being largely complementary. In this way, it is possible to find enough continuity in Wittgenstein's semantic conceptions, which began with the *Notebooks 1914-1916* and ended with *On Certainty*.

[2] The word 'semantics' is understood here in a broad sense that includes pragmatics as the study of words in use, insofar as it is able to influence truth-values of statements.

1. Semantic-cognitive link

In this book, I support the most common viewpoint concerning the referential mechanism, according to which referential expressions can only refer because of some *intermediary link* able to associate them with their reference. This view originated in classical antiquity. A fundamental point to be considered is that this link can always be seen from two contrasting perspectives: the s*emantic* and the *cognitive.*[3] From a cognitive or psychological perspective, the link is usually called an *idea*, *representation*, *intention*, *conception*, *thought*, *belief* and *cognition* (Aristotle and Locke were models of semanticists who have adopted this perspective). From a semantic perspective, the link is more often called *sense*, *meaning*, *use*, *application*, *intension*, *connotation*, *concept*, *informative content*, *belief-content, content of thought, proposition*, *criteria*, *criterial rule, verifiability rule, meaning-rule* (the Stoics, Frege, and Husserl were models of semanticists of this last persuasion).

At this point, an old question arises: What is the appropriate link? Which set of terms should be included or excluded? Should we exclude psychological terms, so as not to contaminate semantics with natural contingency? Or should we abandon a possible commitment to questionable abstract semantic entities, exchanging them for the more feasible concreteness of the psychological, the only thing really able to justify mental causality? Should we read an ambiguous work like the *Critique of Pure Reason* from a semantic or from a psychological perspective?

Traditionally, philosophers have dealt with this problem by assuming that one of these two alternatives must be correct. Nonetheless, this is the real mistake. They have assumed that these two alternatives are mutually exclusive. I see this assumption as a false dilemma, generating useless philosophical disputes. The psychological and semantic perspectives should be seen not as mutually exclusive alternatives, but as complementary and at the bottom inseparable.

The source of the illusion that these two perspectives are irreconcilable lies in the fact that the abstract character of the semantic perspective seems to be committed to some form of realism (Platonist or in a sense

[3] This semantic *versus* cognitive dichotomy can be traced at least as far back as Aristotle. The latter viewed the intermediary link as an affectation of the soul (*ton en têi psychêi pathêmáton*) or thought (*noêmata*) – a psychological perspective – while the Stoics, who appealed to 'what is said' (*lectón*) or 'what is meant' (*semainómenon*), associated the intermediary link in some way with language – a semanticist view. (Manetti 1993: 93 ff.)

Aristotelianist) about universals. In contrast, the cognitivist perspective seems committed to some kind of nominalism or at least particularism attached to the contingency of the psychological subject. Since these ontological commitments are incompatible, the two alternatives also seem incompatible.

However, when we perceive that these ontological commitments could be avoided, it becomes easy to conclude that the intermediary link between words and things can be dealt with in these two apparently contradictory ways without a real conflict. In order to reach this conclusion, we must realize that when we consider the intermediate link from a semantic perspective, we are not necessarily committed to the appeal to the kind of abstract entities assumed by realism. What we are doing is leaving *out of consideration* the inescapable fact that meaning can only exist insofar as it is spatiotemporally embodied in specific psycho-physical subjects (or persons).

In order to clarify the complementarity that I am suggesting, the intermediate link can be considered as both:

(a) a cognitive link, consisting of semantic elements that must be spatiotemporally realized as ephemeral cognitions experienced in specific psycho-physical subjects;
(b) a semantic link, which is referred to as the same semantic elements considered *in abstraction* from their spatiotemporal realization as cognitions going on in a certain specific psychological subject in a specific time and space, but *not* in abstraction from any spatiotemporal instantiation in at least one only particular psycho-physical subject. So considered, the semantic link can be distributed among an indeterminate number of cognitive subjects, even ones not immediately concerned, which does not make it de-psychologized or disembodied (*Cf.* Appendix to this chapter).

In other words: the proposed abstraction cannot be achieved in a sense where the semantic link is considered as somehow transcending the realm of specific psychological and physical subjects, since it always requires some form of cognitive spatiotemporal intentional embodiment in order to be an object of consideration. In fact, the word ‘abstraction’ means here simply leaving out of consideration the natural association between a meaning and this or that specific psycho-physical subject which instantiates the meaning, and focusing on the signs that can convey this meaning, insofar as they can be understood by some other psycho-physical interpreter. This is the only way to make a semantic-cognitive link

semantically independent of its instantiation in occasional cognitive subjects.

A very simple example illustrates my point. When I recognize a patch of vermilion of cinnabar (a precisely characterized shade of color), it is because the patch I see matches a memory image of vermilion that I have stored in my long-term memory from earlier experiences. Now, when I speak of a general concept of vermilion of cinnabar, I intend to show that I am speaking not only of this image, which may become conscious in my mind, but also of *any* other qualitatively identical[4] image of this color that may become conscious in mine or any other mind.[5]

In other words, contrary to the idea that our semantic link is a *type* that is a unique abstract Platonic or Sub-Platonic entity, what I am proposing is that we conceive the semantic link in the sense of an arbitrarily chosen *model*, ideally, *as any token that stands for any other token that is qualitatively identical to it.*[6] In short, we can define a semantic link *X* as:

> A semantic link *X* (*Df.*) = any occurrence of *X* arbitrarily chosen to serve as a model for any other occurrence of some *X* that is qualitatively identical to the model.

Since all these possible occurrences must be psychological (and certainly also physical), we do not need to transcend the domain of the psycho-physical in order to reach the abstract semantic domain. Moreover, we do not need to have an instantiation of the semantic type in any privileged chosen psycho-physical subject. What we really need is for *at least one* psycho-physical subject, no matter which, to embody the semantic cognition. But this condition, as we will see later, can easily be accommodated in our commonsense ontological framework supplied by those particularized properties called tropes.

---

[4] *Qualitative identity* is the identity between different things; it is opposed to *numerical identity*, which is the identity of a thing with itself.

[5] Of course, one could also do the same thing without drawing on color memory: suppose that people carry with them templates of vermilion, so that whenever necessary they compare the patches of color they see with these templates. This shows the importance of some empirically given model, as much as the merely complementary role of memory.

[6] It is true that this last 'any' allows us to infer that there is a class called the class of all tokens that are qualitatively identical, but this class does not belong to the definition and does not need to be an object of awareness.

We can strengthen this compromise solution, if we note that even some sub-items of (a) and (b) show an approximate correspondence to each other. Thus:

(i) the psychological word 'idea' has meaning proximity to the semantic words 'sense' or 'meaning,' as well as to 'concept';
(ii) the psychological word 'representation' has some meaning proximity to the semantic phrase 'criterial rule';
(iii) the psychological phrase 'mental image' has meaning proximity to the semantic phrase 'criterial configuration';
(iv) the psychological word 'belief' has meaning proximity to the semantic phrase 'belief-content.'
(v) the psychological phrase 'occurrence of thought' has meaning proximity to the semantic terms 'content of thought' and 'proposition.'

2. Why reference cannot be meaning

When we consider the semantic link, words that more easily come to mind are 'sense' and 'meaning' (generally used as synonyms), here restricted to cognitive meaning or informational content. However, what is sense or meaning? Perhaps the simplest answer is what might be called *semantic referentialism*, a doctrine that in its crudest form holds that the meaning of a linguistic expression is its own reference. This conception either denies the existence of a semantic link between word and object or minimizes its importance. Wittgenstein described this way of understanding meaning at the beginning of his *Philosophical Investigations*, where he commented on the so-called 'Augustinian conception of language':

> These words, it seems to me, give us a particular picture of the essence of human language. It is this: individual words in language name objects – sentences are combinations of such names. – In this picture of language, we find the roots of the following idea: Every word has a meaning. This meaning is correlated with the word. It is the object for which the word stands. (1984c, part I, sec. 1)

Wittgenstein's aim in this passage was to object to semantic referentialism, a theory championed by him in his first and only published book, the *Tractatus Logico-Philosophicus*. According to his version of semantic referentialism, when completely analyzed, language proves to be composed of atomic

propositions constituted by atomic names whose meanings would be the simple and indestructible objects necessarily referred to by them.[7]

Semantic referentialism is not devoid of intuitive appeal. After all, it is usual to explain the meaning of a concrete substantive by pointing to objects that exemplify what it means. In our childhood, we learned what the word 'chair' means because adults showed us examples of this artifact. And we learn the name of a particular person when this person introduces himself to us with her name. Moreover, we learn what a word means or does not mean respectively through positive and negative examples of its application. All this seems to make credible the idea that meaning may be the object actually referred to. This view has at least an almost palpable simplicity: 'Here is the name "Fido," there is the dog that is its meaning.'[8]

However, there are strong well-known arguments against this naive view of meaning. The most obvious is that you cannot predicate of a meaning what you predicate of an object: if a pickpocket steals your wallet, you do not say that the meaning of your wallet was stolen, and if you say that Julius Caesar was assassinated you do not intend to say that the meaning of his name was assassinated.

Another argument is that many different terms have the same reference, while their senses or meanings are obviously different: the singular terms 'Socrates' and 'the husband of Xantippe' point to the same person, although they clearly have different meanings. And it is worth noting that the opposite seems to be the case with general terms: the predicate '...is fast' in the statement 'Bucephalus is fast' allegedly refers to a particularized property (trope) of Alexander's horse Bucephalus; and the same predicate '…is fast' in the sentence 'Silver is fast' allegedly refers to a particularized property (trope) of another horse, Silver. Although the speed of Bucephalus is numerically different from the speed of Silver, in both sentences the word 'fast' preserves precisely the same meaning.

The most decisive well-known argument against the referentialist view of meaning, however, is the most obvious: it concerns the fact that even when a referential expression has no reference, it does not lose its meaning. The singular term 'Eldorado' and the general term 'phlogiston' do not have any reference, but by no means do they lack a meaning.

For a long time, semantic referentialism has been criticized by natural language philosophers as based on a primitive and misleading understanding

---

[7] As Wittgenstein wrote in the *Tractatus*: 'The name means its object. The object is its meaning.' (1984g, 3.203)

[8] The view was ironized by Gilbert Ryle as the 'Fido-Fido' theory of meaning (1957).

of mechanisms of reference. As John Searle once noted, semantic referentialism 'is a good illustration of the original sin of all metaphysics, the attempt to read real or alleged features of language into the real world' (1969: 164). This might contain some exaggeration, but it isn't wrong.[9]

3. Failure of Russell's atomistic referentialism

Well aware of difficulties like those presented above, Bertrand Russell tried to defend semantic referentialism in a minimalist fashion, taking into account only alleged atomic elements of language and the world. It is instructive to consider his attempt. For Russell, the meaning of at least some foundational terms – called by him *logically proper names* – would have their objects of reference serving as their proper meanings. This could be the case, perhaps, with the word 'red.' After all, as he noted, a blind man is unable to learn the meaning of red, since he is unable to see the color (1994: 194-5; 201-2).[10]

However, it is untenable that the meaning of any word can be given by its reference *tout court*. Changing his example a little, suppose that someone demonstratively applies the word 'vermilion' to an *occurrence* of vermilion of cinnabar, which is a shade of red that in practice the human eye cannot further subdivide (a simpler candidate for 'simple' than Russell's red color, since it does not need to include gradations). Could such an occurrence be the meaning? There is an obvious reason to think that an occurrence of vermilion could not be its meaning: the absence of identity criteria. When we consider the occurrence of vermilion, *it will always be different for each new experience*. This is true if the vermilion is physically considered as an externally given spatiotemporal property, and also true if it is a phenomenal appearance, a *sense-datum*, as Russell preferred. Indeed, if the meaning of 'vermilion' is nothing but a detected occurrence irrespective of its relation to other occurrences of vermilion, then each new occurrence of vermilion should be a new and distinct meaning – an intolerable conclusion!

Russell must have seen this problem, for he found a way to defend his view against it. However, as we will see, it was at the cost of becoming entangled in even worse difficulties. He suggested that the object-meaning of a logically proper name would be something immediately accessible –

---

[9] Metaphysicians of reference have more recently attempted to reassert this primitive form of semantic referentialism (*Cf.* Salmon 1993).

[10] As Russell recognized, logical atomism was first suggested by Wittgenstein, who defended it in a full-fledged way in his *Tractatus Logico-Philosophicus*.

such as *sense-data* picked out by pronouns like 'this' or 'that' – only as long as we keep these sense-data present in our consciousness… This means that the meaning also lasts only as long as our personal experience of a word's object of application! (Russell 1994: 201, 203) However, this is a desperate answer, as clearly it leads to solipsism.[11] What criteria of correction could we apply to fix this 'meaning,' in order to know in what cases the word 'vermilion' can be reapplied to other occurrences of the sense-data that would at least qualitatively be *the same* sense-data? Moreover, how could these logically proper names have the intended foundational role in a language? How could we insert this fugitive meaning of a proper name in our common language – a language composed of words whose meanings are permanently shared by their speakers?

We need to acknowledge that in our language, to know the meaning of a word like 'vermilion' presupposes at least the ability to recognize an occurrence of vermilion as being precisely similar to other occurrences of vermilion. But this acknowledgment is not included in the idea that the meaning of a word is nothing more than the occurrence of its reference. The concept of a word's meaning essentially requires that we should be able to *unify* its different applications to the same referent, which is not possible by means of Russell's account alone.

It is true that if the meaning of a word like 'vermilion' were the vermilion-*type* – understood as an abstract entity common to all occurrences (*tokens*) – we would be able to solve the difficulty pointed out above. But this solution might commit us to accepting some form of (Platonic or sub-Platonic) realism, raising justified suspicion of an unintelligible reification of the *type* in a *topos atopos*. Alternatively, one would need to consider the vermilion-*type* as being a certain set of occurrences of *sense-data* that are precisely similar to each other. This reduces the danger of realism, but does not eliminate it, since sets are often seen as abstract entities, and if they are not, then they need here some limiting intention. In addition, sets may be larger or smaller depending on how many members they have, while the meaning of the word 'vermilion' has no proper magnitude, neither increasing nor decreasing. Even the resource to open sets would not be helpful since they are abstract constructs and not what is effectively given.

[11] This kind of difficulty was already raised in the public discussion of Russell's speech in 'The Philosophy of Logical Atomism,' 1994: 203. (For criticisms see Tugendhat 1976: 382, and Kripke 2013: 15-16.)

The most feasible alternative seems to be that we consider the meaning of 'vermilion' as some occurrence of vermilion that we are using as a *model.* This could be either a *sense-datum* or some particularized property in the outside world, able if necessary to be arbitrarily replaced by another like it or any other occurrence that is precisely similar to this model. Thus, if I recognize what is currently being offered as an occurrence of vermilion, it may be because I realize that this occurrence is qualitatively similar to others that were previously given to me as being those of vermilion. This relies on a model whose copy I have stored in my memory, giving me an awareness of it as a color qualitatively identical to colors I have previously experienced. Thus, recalling the various experienced occurrences of vermilion {V1, V2... Vn} I must have a model Vm in my memory. Hence, I can say that V1 = Vm, V2 = Vm... Vn = Vm and, therefore, that V1 = Vm = Vn, etc. I can do this without resorting to any Platonic entity or to any multiplication of identities of identities or even to the concept of an intentionally defined set – problems often thought to burden particularistic strategies for handling universals.

What this view amounts to is that what we could call the referential meaning of the word 'vermilion' must be identified with a *referential connection* (a true relation of remembered similarity). Now, this referential connection is a *rule* that relates cognitive experiences of occurrences of a color to occurrences of color that we in some way use as models, in order to produce an awareness of what is experienced as being qualitatively identical vermilion colors in each case. Moreover, this internal semantic cognition is produced in association with 'vermilion': the concept-word for such entities. In this way, both a reference and its name turn out to be in principle interpersonally accessible, once the qualitative identity between occurrences associated with the same word allows for interpersonal accessibility and for the kind of practical implicit agreement necessary to create a linguistic convention. Indeed, this convention can be created, even if in itself the semantic cognition isn't, as a matter of fact, interpersonally accessible.[12] We should also point out that the *semantic rule* that uses recollections of models to identify any new instance of vermilion is independent of this or that particular occurrence of vermilion, for it only relates to instantiations of possible occurrences that can satisfy it. This kind of solution is the only that seems to be workable.

However, this solution has a price: we see on reflection that by adopting it we have already left behind the referentialist conception of

[12] See my discussion of Wittgenstein's private language argument at the end of this chapter.

meaning. Even to establish a meaning as simple as that expressed by the word 'vermilion of cinnabar' we must appeal to something that is more than a rough object of reference and is independent of it, namely, a *semantic rule*.

Even if Russell's semantic referentialism is unsustainable, there is a lesson to be learned from discussing it. Our last suggestion recovers an important idea derived from his semantic referentialism, namely, the idea that the existence of an object of reference is *necessary* for the names of objects taken as simple in the context of a certain linguistic practice. It is true that we always need to understand the phrase 'simple object' as inevitably having a non-absolute sense restricting it to a non-decomposable entity in the framework of some linguistic practice (Wittgenstein 1984c, I, sec. 45-48). And this would be the case with the *sense-datum* or external property of vermilion of cinnabar as a *trope* (a spatiotemporally particularized property). The positive conclusion is that *for such 'simple names' to acquire appropriate meaning they need to have reference.*

This is why, in an important sense, a blind man cannot learn the meaning of the word 'red.' Since the color red is in a sense simple[13] and knowledge of it is based on acquaintance, and since the blind cannot have this sensory experience, a blind man cannot learn and apply the conventional criterial rule for the shared referential meaning of the word 'red.' At least in the case of this subrogate of a logically proper name restricted to a certain linguistic practice, the existence of some object of reference is indispensable. But this obviously does not lead to the idea that a word's reference is its meaning. What it means is only that in basic cases a given object of reference is indispensable for the formation of the semantic rule whereby a word acquires its referential semantic function.

[13] One could object that since there are many different shades of red (one of them being vermilion), red cannot be simple. But with Wittgenstein we can answer that what we call 'simple' depends on whatever linguistic system we have adopted: we can use an old linguistic practice with only three basic colors: red, yellow and blue. Here red will be considered simple; and in this case, distinct shades of red will not be taken into account, even if they are perceptually distinguishable. Instead of being qualitatively identical to the pattern, a new red patch must only be *sufficiently identical*, insofar as we have parameters to distinguish it from the blurred borders with the other two colors.

4. Meaning as a function of use

We now move on to a second candidate for the semantic link: use or application. Wittgenstein privileged this candidate, suggesting that the meaning of a linguistic expression is its use (*Gebrauch*) or application (*Verwendung*). As he wrote in a famous passage of *Philosophical Investigations*:

> You can, for a large class of cases of use of the word 'meaning' – if not for all cases of its use –, explain it like this: the meaning of a word is its use in a language. (1984c, part I, sec. 43)

This suggestion applies to both words and sentences. It clearly applies to (a) what has been called *directive meaning*: the illocutionary forces of expressions, which establish kinds of interaction between speaker and hearer in speech acts and can be made explicit by so-called performative verbs like 'I order,' 'I promise,' 'I quit'… However, directive meaning, together with (b) *expressive meaning*, which aims to express internal psychological states, though also considered by Wittgenstein, are two kinds of meaning with little relevance for us here. The focus of our research is the kind of meaning able to link our linguistic expressions with the world, something that is sometimes called (c) the *referential meaning* – the kind of meaning typically required for the communication of information.

My concern here, as was clear right from the start, is the *semantic content* of declarative sentences, which is the kind of referential meaning we call *cognitive, epistemic, informative, descriptive* or *factual,* able to link language with the world and to be endowed with truth-value (Aristotle's *logos apophantikós*). Such epistemic, informative or descriptive semantic content should be of major philosophical importance, because by being able to relate language to the world, it should have epistemological and ontological import.

However, the identification of meaning with use doesn't apply so easily to the cognitive or referential meanings of our sentences and terms. Consider, for example, a declarative sentence like 'The tide is high.' It is easy to imagine an illocutionary use for this sentence, such as warning or informing. However, by identifying meaning with in such cases we would revert to meaning as force. In his theory of speech acts John Searle has distinguished in all utterances the necessary form $F(p)$, where (explicitly or not) F expresses an illocutionary force and p (explicitly or not) expresses a propositional content (1983: 6); no speech act makes real sense without the combination of these two elements. Anyway, if we wish

to approach use with cognitive meaning, with and without force we must attend to the use of *p* as *p*, which is not the easiest thing to do.

It is possible to approach pure cognitive or referential meaning with an appeal to use by producing an acceptable *extension* of the concept of use. Consider first the cognitive meaning of *p* as *p* without judicative and assertoric force. We can isolate cognitive meaning from force by employing the Fregean device of expressing a sentence's content only as being *regarded*, depriving it of any assertoric force. We can do this by making a sentence like 'The dog has run away' the subordinate clause of 'It is possible *that the dog has run away*.' The spelling of the complementary sentence '…that the dog has run away', expressing cognitive content, even if not asserted – could also be seen as *a use.* And *use* could also in this case be considered to be any realization of a phonetic shape of the mental construction of a verifiability rule constitutive of the sense/meaning of the subordinate clause as conceivably (though not as really) applicable – which in fact does not identify use with meaning.

But we can also try to approach use to the cognitive or referential meaning involved in the whole act of communication by means of which a speaker intends to share with a hearer his awareness of a real or possible fact. For instance: when a speaker says 'The tide is high,' the use may involve (i) *the utterance in which a propositional content* (cognitive meaning) *is expressed,* added to (ii) the *assertoric force as an external expression of the judicative force*. Here the speaker intends to communicatively reproduce the same judgment (the same propositional content plus its judicative force) in the hearer's mind. In an extended sense this can also be called use: this is use as *communication of the judication of a cognitive meaning*, the last being what one might suppose to be a verifiability rule applied to a real-actual fact (*Cf.* Chapters IV and V of this book). If not the identity, we see the narrowness.

But what about the hearer's understanding of a statement? The hearer is not using phonetic shapes in his understanding of its meaning. In order to maintain the view that even in this case meaning can be approached to use, we need to resort to a bolder extension of the word. It seems possible to say that we use expressions referentially or not, simply by *thinking what we mean when we spell them*. When a hearer really thinks the tide is high, it is possible to say that he actually uses this sentence in an epistemic mode by thinking it. Thus, if Paul understands the sentence 'The tide is high', or if Anne comes to believe that 'the dog has run away,' with or without using words, Paul is repeating (or interpreting) and Anne is producing the judgments of these respective contents internally, that is, they are applying the supposed verifiability rules of these sentences merely

in thought. Hence, in normal communication, the use that a hearer gives to heard words by understanding them could consist in conceiving the construction of verifiability rules with their identification and ascription rules in a way similar to what the speaker should do when using words to convey cognitive meaning. The conclusion is that not only the cognitive meaning as the speaker's thought, but also the hearer's thought, could be viewed as *internalized cognitive way of use*, with or without the addition of judicative force, which could also be seen as an internalized form of assertoric force. Finally, if Plato was right that discursive thought is 'a silent dialogue of the soul with itself,' we can generalize this process of internalization and consider any cognitive act as a way of use, even without being associated with communicative action.[14] Associating it with language, we might call this the *cognitive use* of an expression, of which judicative and assertoric forces are only complementary elements.[15]

It is easy to find objections to the relevance of the proposed extensions of the meaning of the word 'use' that I am employing in order to save a supposed identification of use with meaning. Indeed, though they do not seem to be wrong, they can be considered too confusing and cumbersome to justify themselves. However, as will become clear, the real reason why Wittgenstein viewed meaning as a function of use was a different one. It was the pragmatic advantage of locating meaning in its most proper place from the start: in normal linguistic praxis, in the concrete speech-act situation, even in the normal practice of thinking with words. This enables us to individuate the meaning of an expression where it exercises its proper function, so that in this way we achieve the highest level of

[14] Language not only has a *communicational* function, but also an *organizational* function, in the sense that we also use it to think, to organize our ideas and our plans of action (Vygotsky). At first sight, the identification of meaning with ways of use doesn't seem to do justice to its organizational function, but this doesn't have to be so. It makes sense to say that when I think that the Leaning Tower of Pisa could come crashing down, I am *using* this name referentially in my mind, in thought, that is, in an internal dialogue with myself.

[15] In insisting that the content of p is a communicable kind of meaning, I distinguish this analysis from the Gricean psychological theory of meaning. H. P. Grice suggested that to display what he calls a *non-natural meaning* (our semantic-cognitive meaning) of *p* the speaker must have the intention (i) that the hearer should come to believe that *p*, (ii) that the hearer should recognize the intention (i) of the speaker, and (iii) that by means of the recognition of (ii), the hearer will come to believe that *p*. However, what Grice thereby analyzes is not the non-natural meaning in itself, but only the standard procedure by which the non-natural meaning is communicated. (*Cf.* Grice 1991; see also Tugendhat 1976, Ch. 14).

contextual and interpersonal corrigibility, with a minimal amount of distortion and exclusion.

This is, I believe, what Wittgenstein's identification of meaning with use is all about: It allows us to individuate meanings in the natural contexts of their existence, while in doing philosophy we are too easily prone to decontextualize meanings, excerpting and distorting them, in order to develop insights that can be highly illusory. In this sense the maxim that meaning is a function of use can help us in practicing what Wittgenstein called philosophy as therapy, which aims to untie the knots of thought tied by philosophers, insofar as it brings our words back from their metaphysical holidays to their daily chores (Wittgenstein 1984c, part I, sec 116).

5. Meaning as a kind of rule

A related point arises when we perceive that a really appropriate identification of meaning with use cannot be one of meaning and *episodic use tout court*, namely, a mere spatiotemporal *occurrence* (token) of a linguistic expression. This isn't possible, because each occurrence differs from others in its spatiotemporal location. If it were the case, each new occurrence would be a new meaning, which would result in the semantic catastrophe of making the number of meanings of any linguistic expression unlimited.

There is, however, a more reasonable alternative. We can understand the words 'use' (*Gebrauch*) or 'application' (*Verwendung*) as an abbreviation of *way of use* (*Gebrauchsweise*) or *way of application* (*Verwendungsweise*), since the same word can be used many times in the same way. But what is the way of use? Well, it doesn't seem to be anything other than 'something of-the-type-of-a-rule' (*etwas Regelartiges*) that determines episodic uses. Wittgenstein himself came to that conclusion in an important, though less well-known passage of his last work, *On Certainty*:

> The meaning of a word is its *mode* of application (*Art der Verwendung*) ... Hence, there is a correspondence between the concepts of 'meaning' and 'rule.' (1984a, sec. 61-62)

In fact, to use a word meaningfully is to use it in accordance with its mode or way of use or application, it is to use it correctly, and to use an expression correctly, in the right way, is to use it in accordance with those rules that give it its meaning. By analogy, we can say that we use a screwdriver according to its way of use when we use it correctly, according to a rule, turning it clockwise in order to tighten a screw.

Consider the following examples of ways of use based on the *Linguee Online Dictionary*, which includes numerous examples of words used in sentences:

> WAY OF USE: Apply several times to the skin and rub in for several minutes with a circular motion, until completely absorbed.
> WAY OF USE: To color and cover up grey hair, we recommend 20 ml. 6% of a cream oxidizing agent in the proportion of 1 + 1.
> WAY OF USE: Never dispense any pharmaceutical product without a prescription detailing way of use, site, withdrawal periods and other relevant information signed by a physician.

Of course, here 'way of use' means rules or sequences or combinations of rules for correctly using things. Now we see clearly that meaning can only be identified with use in the sense of 'something of-the-type-of-a-rule' determining episodic uses. And what holds in general for a word's use also holds here for cognitive or referential use.

In fact, the identification between meaningfulness and rule is more primitive. Consider the following two signs: **'OO'** and '**O→**.' The second seems to us 'more meaningful,' since we have the tendency to link it with a rule pointing in a particular direction. Rules are the ultimate intrinsic source of meaningfulness.

6. Meaning as combinations of rules

However, why does Wittgenstein prefer to say that meaning is *determined* by rules? Why cannot the meaning of our linguistic expressions be identified with rules *simpliciter*?

In my view, at least part of the answer was also approached by him with his analogy between language and *calculation*. (Wittgenstein 1984f: 168; 1982: 96-97). This understanding is reinforced by the many otherwise unjustified considerations in his *Brown Book* of how complex sequences of rules could be followed in relatively simple language-games, understood as systems of rules. In use, linguistic expressions normally involve calculations, which should be understood as nothing more than *combinations of conventional rules*. And the cognitive meanings that these expressions have can consist essentially in combinations of more or less implicit, automatized semantic conventions, knowledge of which speakers tacitly share.

Arithmetic can serve as an illustration. If the meaning of a mathematical proposition is constituted by its proof, considering that proof is a

combination of rules, this meaning is also a combination of rules. Some people can do the multiplication '120 x 30 = 3,600,' for instance, by combining three rules, first multiplying 100 by 30, then multiplying 30 by 20, and finally adding the results 3,000 and 600 to get the result 3,600. The meaning understood as the cognitive content of multiplying '120 *x* 30 = 3,600,' would be given by this and other methods of calculation. Together they should amount to essentially the same general signification – what I would call the same *rule-complex* (*Regelkomplex*) – insofar as they proceed in different but complementary ways, i.e., beginning at the same starting points and reaching the same final result.

We see that what we called 'something of-the-type-of-a-rule' can be understood as possible combinations of rules that starting from some initial conditions bring us to some final result. The cognitive meaning of a linguistic expression must also be the same as (i) *a specific semantic-cognitive rule* or (ii) one or more *combinations of semantic-cognitive rules* that determine a correct episodic use of the rules. And the cognitive meaning of a linguistic expression is a rule, combination of rules, or a rule-complex that when applied or satisfied brings about a cognition of some state of affairs. Calling such combinations *rules* – as I do in the present book – is ultimately a justified extension of the term 'rule,' since combinations of rules and a variety of combinations of rules that produce the same final results do the same job as rules. Although irreducible to implicitly shared conventions, such combinations can still be seen as *conventionally grounded*, since they are constituted by elementary rules, namely, ones usually established by convention. Thus, when someone says, as Wittgenstein sometimes does, that meaning is *determined* by rules, what can be reasonably understood by this is that cognitive meaning may be the application of some combination of rules or some variable combination building a rule-complex enabling us to reach the same cognitive effect, and nothing more.

Since we are interested in the problem of reference, the meaning that will be considered will be a content – called cognitive, informative, epistemic or referential – that is, something reducible to semantic-cognitive rules responsible for our linguistic awareness of what can be objectively given, which are also *criterial rules*. So, we are dealing with cognitive-criterial rules responsible for the cognitive or referential significance of declarative sentences. Criteria are, in Wittgenstein's own terms, 'what confers to our words their ordinary meanings,' (1958: 57). As I understand him, semantic-cognitive rules are based on criteria or criterial configurations, which are conditions generated by these rules, insofar parts of them, and hence part of their meaning-giving function. On the other hand, criteria

(having process-product ambiguity) can also be considered those cognitively independent conditions that once given satisfy such dependent or internal criterial conditions produced by the criterial rule, making us realize that something is the case. Using Wittgenstein's own example, if someone says 'It's raining' and this statement is true, this involves applying a criterial rule, a rule which requires that certain conditions must be given – say, drops of water falling from the sky – so that a cognitive awareness that it is raining follows (2001: 28). And this resulting awareness, the cognition, could be understood, as already suggested at the end of the first chapter, as the availability to the system of what results from criterial conditions definitely accepted as satisfied.

However, if an analysis of the appeal to use leads us to cognitive reference-rules, why appeal to use? Why not just start with an investigation of these rules and their combinations? The answer was already given. Language is primarily an instrument of action, and meaning, cognitive or not, is there mainly to facilitate action. Moreover, semantic rules are not abstract objects in any realist sense, since this is an old philosophical trap (See Appendix to this chapter). If not dispositionally considered, these rules only exist when they are applied. Therefore, attention to correct use helps us to individuate meaning and to find the real cognitive-criterial rules or combinations of rules that must of necessity be applicable, that is, applied either in reality or in imagination, in order to confer meaning to a fully contextualized linguistic expression.

We can further elucidate what is at stake by appealing to a metaphor: when a post office delivers a letter, the envelope gives general indications as to the addressee's geographic location (city, state, country, etc.). These general indications can be compared with the grammatical meaning of a sentence and also with its logical analysis. But even if necessary, they are not sufficient. Too many other addressees live in the same country or city or street, just as too many different sentences have the same grammatical or even logical structure. To reduce this vagueness, mail carriers also need the name of the street, the building or house number… Without singularizing details, it can be almost impossible to deliver mail to its proper destination. The same holds for cognitive meaning. It can be decisive to care about the way of applying our expressions in a given context, which can be the whole discursive and practical context, including that of philosophical writings. What an appeal to use does is to lead us to semantic details. The most general traits of an expression's way of use, though relevant, are common to many other expressions and for this reason in themselves insufficient to individuate meanings. Because of this, the more specific traits of meaning specified in ways of use are also

important. And these are traits that expressions can only gain in the real contexts where they are applied. Consequently, these can only be fully explored by surveilling linguistic praxis. This is why it is so important to explicitly consider occasions of use in all their pertinent details. Indeed, the main flaw of many philosophical examples and thought-experiments consists in ignoring apparently irrelevant subtleties. These can be responsible for easily ignored subtle semantic variants that an expression can have in different particular contexts. Consideration of such subtle semantic differences is of particular importance for correcting misconceptions arising from philosophical attempts to use words beyond the limits of meaningful language. Particularly elaborated philosophical examples of overstepping these limits are those concerning the metaphysics of reference and meaning.

7. Meaning and language-games

There is more to be said about meaning as a function of use. The first thing to note is that a linguistic expression only makes sense when used within a system of rules often called a language-game. To explain this we might again appeal to a metaphor. We can compare a linguistic expression with a chess piece, and its use with a move in playing chess. When we move a chess piece, the meaning of the move is minimally given by the rule that governs the piece's move. But what the move fundamentally means will depend on the game situation. It will be given by the contextually determined tactic, by the calculation of possible combinations of rules in anticipation of possible moves by the opponent and responses that could be made.

Something not very dissimilar occurs with linguistic use. The linguistic rules governing what Wittgenstein called 'superficial grammar' could be compared to the rules for moving chess pieces. But these grammatical rules – even those of logical grammar – may not be what really matters. Often what is essential are rules, rule-combinations and rule-complexes belonging to what he called 'deep grammar' (1984c I, sec. 668). These may have more resemblance to semantic-cognitive rules like those we exemplified before (for the proper name 'Aristotle' and for the concept-word 'chair'). Their combination would justify moves that suggest chess players' tactical calculations, which is particularly clear when we consider dialogical speech.

To give an example. One knows that the sentence 'Calphurnia urged Caesar to stay at home' is grammatically correct, and one may even know that its logical form is *aRb*. But this will be of no help if one does not

know who Caesar and Calphurnia were, where and when they lived, what relationship they had, and cannot even imagine when or why she has warned him to stay at home. Superficial grammar (or syntax) gives expression to a grammatical sense that is often the same for semantically different sentences. But the semantically relevant rules and combinations of rules that constitute what is meant by a linguistic expression are more flexible and might change not only with the sentence, but also in accordance with the particular factual and linguistic-discursive contexts.

Furthermore, in a similar way as the rules-combinations responsible for a strategic move in chess gain their meaningfulness depending on the changeable state-context provided by the system of rules that constitutes the game of chess. And the rules determining the application of linguistic expressions are able to produce meaningful utterances only when combined in the changeable context furnished by the system of rules constitutive of the language-game.

Anyway, a chess metaphor is too liberal, insofar as it does not take account of what Wittgenstein would call the divisions of language. What he called a *language-game* (*Sprachspiel*) or a *linguistic practice* (*sprachliche Praxis*) is more adequate. A language game can be understood as *any linguistic system of rules that typically includes syntactic, semantic and pragmatic rules that belong to our language*.[16] Examples of language-games given by Wittgenstein are:

> Giving orders and acting according to them, describing an object by its appearance or measures, informing… speculating about an event, making and testing hypotheses… making up a story, reading… solving a riddle, telling a joke, describing a landscape, acting, asking, thanking, cursing, greeting, praying, etc. (1984c, sec, 23)

But he also uses the same idea in a wider sense, pointing to more extended domains of language like:

> The language of colors, the language of proper names, or even the important 'knowing games' from *On Certainty*, like the game of doubt and the languages of history, physics, chemistry, and arithmetic. (*Cf.* Costa 1990: 50)

That is: it seems that almost any semiotic chunk of our language, insofar as it is identifiable as such, can be seen as a language-game. Language-

[16] There are also experimental, simplified or artificial language-games that the philosopher invents to make comparisons… But I am interested here in the language games really constitutive of our natural language.

games *include* themselves, one within another, like the case of Cantor's theory of infinite numbers within the theory of numbers and the language of mathematics; and they can *partially overlap* one another, as when someone describes a scenario and simultaneously tells a joke, insofar as we remain able to distinguish them (Wittgenstein 1984c, sec. 46-48). Fundamental is that the language-games remain identifiable at the interpersonal level.

The concept of language-game or linguistic practice contains the concept of the speech act, systematically studied by J. L. Austin and John Searle, but it is much wider. This is why Wittgenstein was not mistaken when he wrote that there are countless language-games (1984c I, sec. 23).

By making the meanings of expressions the result of rule combinations belonging to rule-systems typified by language-games, Wittgenstein was endorsing what was later called *semantic molecularism*: What we call the meaning of an expression does not depend on the expression in isolation (semantic atomism), nor on its insertion in language as a whole (semantic holism). It depends more properly on the often variable *state-contexts* of the linguistic practice in which it is inserted (a molecular subsystem of language). Finally, it is a mistake to believe that meaning is a matter of all or nothing. It is much more reasonable to think that when used according to the rules of a language-game, something of a word's meaning gradually merges into a maze of partially related meaning-rules.

In support of the idea that we use and give meaning to the expressions of our own language in language-games, in his *Brown Book* Wittgenstein described natural language as a great nebula of language-games:

> The language of the adult presents itself to our eyes as a massive nebula, natural language, surrounded by more or less defined language-games, which are technical languages. (1984e: 122)

Later, in his *Philosophical Investigations*, he compared language to a great old city:

> Our language can be seen as an ancient city: a maze of little streets and squares, of old and new houses, and of houses with additions from various periods; and this is surrounded by a multitude of new boroughs, with straight, regular streets and uniform houses. (1984c, sec. 18)

The nebula, the city, begins with what was built in its original center: the practices of ordinary language, expressing our ordinary commonsense wisdom. To this, there come new insights, like those better organized language-games arising with the emergence of new scientific fields. As

with games, the great old city can be subdivided in many distinct ways, one part including another and one part overlapping another.

There is a noteworthy relation of dependence here: learning and teaching new scientific and technical practices, even the possibility of their understanding and creation, depends on the prior acquisition of more basic language-games governing ordinary life. This coheres with our principle of the primacy of established knowledge (Ch. II. Sec. 4), which leads us to conclude that rejecting the assumptions of our *modest* common sense by means of science would be a very questionable approach.

A question that now arises is: in such circumstances, what criteria would we have for identifying meaning variations, or, less ambitiously, what criteria would we have for identifying the language-game in which an expression is used or even misused? Considering that language can be subdivided into multiple and varied ways, it seems that we can apply different criteria to the same linguistic move, insofar as we are able to interpersonally identify and share the criteria we are applying... But in this case, what guides us in choosing a criterion? Is this identification really possible?!

I believe that an affirmative answer is possible. My tentative suggestion is that the identification of a language-game according to the criteria for the use of an expression (term, phrase, sentence), which also establishes the shareable meaning of the expression, involves what we could call *identifying state-contexts*, which are created by two factors:

(i) the *relevant factual and linguistic context determining the expression's use*, together with
(ii) the *speaker's intention* in using the word, insofar as this intention is made interpersonally clear, either by spelling or in a contextualized tacit way.

It seems that in the normal case awareness of these two factors, namely, of the state-context of the words' application by the speaker is what allows the public identification of the relevant language-game in which he is using a linguistic expression and in this way, the relevant meaning rules meant by him. On the other hand, it seems that if the hearer correctly identifies the speaker's state-context – the right given context implying the intention and possibly complemented by the spoken intention – he identifies the language-game the speaker has in mind and will be able to understand correctly what the speaker means. (A simple case: if a teacher told his students that the philosopher who represented the culmination of the philosophical thought of antiquity was called 'Aristotle,' the context

shows everyone that he was playing a game of naming in which he intended to speak about the famous Greek philosopher and not about someone else with the same name, despite the fact that this game of naming is included in a game of teaching, which is included in the game of public speaking.)

8. Meaning and form of life

There is a last important concept in the understanding of Wittgenstein's explanation of meaning. The linguistic practices that form the nebula find their ultimate *raison d'être* as constituents of what Wittgenstein called a *form of life* (*Lebensform*). As he wrote in his few passages on this concept:

> …the word 'language-game' is used here to emphasize the fact that *speaking* a language is part of an activity, or of a form of life. (1984c, I, sec. 23)
>
> Right or false is what human beings say; and in the language they agree on. This is no agreement in opinions, but in form of life. (1984c, I, sec. 241)
>
> What is taken for granted, the given, we could say, are forms of life (1984c, II: 572)

He arrived at this foundational idea probably influenced by an article written by the great anthropologist Bronislaw Malinowski, who suggested that in order to learn the language of a primitive people one needs to share life with them in their society (Malinowski 1989).[17] One example used by Malinowski to illustrate this point can be revealing here: when fishermen in the Trobriand Islands use the phrase 'paddling in a place,' they mean they are navigating close to an island village. The waters around the islands are so deep that it is not possible to use a pole to propel a canoe, so they need to paddle their boats to reach the village. Only by knowing speakers' life circumstances can we find the information needed to understand what their expressions mean.

The relevance of much that Wittgenstein wrote consists in his having seen the importance and comprehensiveness of some ideas. For him, the phrase 'form of life' means *the way of life in a society*. More precisely: *the*

[17] Although Wittgenstein expressly disliked K. Ogden and I. A. Richards' book 'The Meaning of Meaning', he must have appreciated the short supplement to the book in which Malinowski presents these ideas.

complex of regularities that govern the lives of people in the totality of their social and physical environment.

We can compare the idea of a form of life with what is involved in two technical terms introduced by J. R. Searle. These are (a) the *network* of meanings involved in the determination of an *intention*, and (b) the *background* of abilities, skills, dispositions, and ways of doing things that are linked with the corresponding network (Searle 1983, Ch. 5). Though including what Searle means by network and background, the concept of form of life is more comprehensive, since even the landscape in which a tribe lives should be comprehended by the concept and may have some influence on the meaning.

More auspicious is a comparison between the concept of form of life and Husserl's concept of *life-world* (*Lebenswelt*), which for the latter author can be *the whole of our shared communal world of human activity* (Husserl 1954, Vol. VI: 105 f.), grounding in this way all possible knowledge. For Husserl the life-world, which can be subdivided into a multiplicity of different *home-worlds* (*Heimwelten*), forms the holistic framework within which all knowledge is acquired, serving therefore as the ultimate foundation of all human cultural endeavors, gradually extending into scientific ones. Furthermore, although there are different life-worlds, they must have grounding commonalities: aspects like spatiotemporality, materiality, life, birth, death, instincts, hunger, thirst, etc.

Wittgenstein would probably share this view, at least in its non-theoretical aspects. The comparison shows us something important: we now see that there must be something common in the most basic levels of our different forms of life. For there must be the share of grounding commonalities that serves as a condition enabling us to accommodate ourselves to different forms of life and be able to learn and incorporate other cultures' languages. What enables us to do this is certainly that we all share a fundamentally common human nature and a similar surrounding world.

9. Tying the threads together

We can now summarize. Language appears in Wittgenstein's philosophy as an immensely complex system of syntactic, semantic and pragmatic rules: a system we can subdivide in many ways into subsystems called languages, sub-languages and language-games or linguistic practices, which are in turn rooted in a wider ground: the life-form made up of regularities that determine the lives of people in social groups. Linguistic

practices constituting our natural language originate spontaneously from our form of life and depend upon it. Here again, we see that creating and learning the specialized language-games of science is only possible because of the assumption of more central practices of natural language ultimately entrenched in life-forms. This is also why an inorganic computer will never be able to give meaning to the signs with which it operates: a silicon-based machine is a by-product manufactured by a life-form and not a biological agent naturally growing within it.

We can summarize Wittgenstein view on meaning in a formula:

> The meaning given to an episodic use of the expression X (*Df.*): the compliance of this use with rules in the context of an appropriate linguistic practice (the language-game) rooted in a form of life.

This is a characterization of meaning as something that belongs to the praxis of language as it is understood and to our extensions of the concept of use as what is cognitively meant. This assimilation of cognitive meaning to action by means of an extended notion of use as the mode of use and as a rule-in-its-application is what makes it unnecessary to hypostasize semantic rules as abstract objects in any Platonist sense. Meaning is what we think of or speak about as being meaningful; and what we think or speak is meaningful insofar as it is correctly used, namely, used in accordance with the meaning-rules of linguistic practices rooted in our life-form; and the most relevant meaning-rules are the semantic-cognitive ones, allowing us to represent the world.

This is what I believe we can achieve, based on Wittgenstein's semantic views: an uncomfortably vague but sufficiently plausible and, I think, minimally distorted surveillable representation of the deep grammar of the concept of meaning. This kind of representation is important insofar as it plays a role as a semantic foundation for philosophy as therapy. This is also why a surveillable representation of the grammar of meaning is central to Wittgenstein's thought: it is the sustaining core of his philosophy, as much as the doctrine of ideas was the sustaining core of Plato's philosophy.

10. Criteria and symptoms revisited

Another important distinction that we owe to Wittgenstein, already introduced in Chapter II of this book (sec. 8), is the distinction between *criteria* and *symptoms*. Semantic-cognitive rules are criterial rules. Criterial rules are ones based on conditions called criteria.

There is, as we have also noted, a fundamental difference between criteria and symptoms. Criteria are conventionally grounded conditions that, once accepted as really given, *warrant* for us the application of a semantic-cognitive rule. Symptoms, on the other hand, are conditions that, once accepted as really given, make the application of a semantic-cognitive rule only more or less probable. A criterion should establish the sufficient conditions for the application of an expression, though not properly as given essences, insofar as criteria for the same rule can be often multiple and varied, as our investigation of proper names has shown. Because of this sufficiency, Wittgenstein also called them *definitional criteria*, since their description is definitional of an expression or at least takes part in its definition. They are *primary* criteria, while symptoms are also called *secondary* criteria (*Cf.* 2001: 28).

One example makes Wittgenstein's distinction clear: a criterion for the application of the concept-word 'malaria' is actually finding a bacterium – Plasmodium falciparum – in a patient's blood. Once we assume that we have found this, by definition we are warranted in saying that the patient has malaria. But if all we find is that the person has a cyclically high fever, we have only a symptom of malaria, perhaps a secondary criterion, something that makes it probable that the patient has contracted the disease.

Insofar as criteria are also understood as internal constitutive conditions of the semantic-cognitive rules for the referential use of a conceptual expression (Ch. II, sec. 8), they must belong to its *meaning*, since these rules (whether effectively applied or only imaginatively regarded in their possible application) are constitutive of meaning. When Wittgenstein wrote that criteria 'give words their common meanings' (1975: 57), he was referring to criterial rules.

Finally, criteria play the role of criteria only in the context of the language-games to which they belong. This is the main reason why Wittgenstein says that there can be a grammatical oscillation between criteria and symptoms. With the alternation of linguistic practice, criteria can become symptoms and vice versa (1983c, sec. 79, 354). That is: the same condition that works as a criterion in one practice can serve only as a symptom in another practice and vice versa. And similar changes can also occur as a result of the evolution of language, which may change and improve our conventions, often turning criteria into symptoms by replacing them with new conditions.

The distinction between criteria and symptoms is also important for the critique of language. Philosophers are all too often inclined to treat symptoms as though they were criteria. To give a very trivial example:

consider peoples' facial and bodily features. These are the physical characteristics by means of which we are able to immediately identify people we know. At first sight, it seems that they are the real criteria for identifying persons – and within some superficial language-games they may work in this way. But if we look more closely, we clearly see that they aren't. If a person, as happens in fairy tales, were transformed into a donkey, but continued to behave no differently than before, talking to us and in full possession of his memories, personality, knowledge, and abilities, we would be forced to admit that he remained the same person, even though in a different body. This and other more plausible thought-experiments show that people's facial and bodily appearances are not primary criteria at all, but only symptoms able to make their personal identification probable in an easy and immediate way. To find the ultimate criteria of personal identity is still today a controversial philosophical problem.[18] However, physical appearances will be treated as criteria in the context of some practical language-game, like that of taking attendance in a school class.

11. Transgressions of the internal limits of language

In the *Tractatus,* Wittgenstein was interested in ascertaining what David Pears has called the *external* limits of language and its transgressions (1970, Ch. 5). This is relatively easy to spot: a logical contradiction is an external transgression. However, he came to see that most philosophical confusions are caused by the more subtle transgressions of the *internal* limits of language. These transgressions happen because many of our expressions can be used in different linguistic practices, undergoing in this way more or less subtle changes in meaning. As Wittgenstein also wrote, 'The place of a word in grammar is its meaning' (1984d, sec. 23), a place that cannot be fixed beforehand, since it may circumstantially change. Now, when an expression is used simultaneously in different practices, where it should receive a different meaning or meaning-modulation, it turns out to be easier to confuse what we mean with it.

In Wittgenstein's philosophy, we can find two forms of confusion or misleading uses of expressions, which we may call *equivocity* and *hypostasis*.[19]

[18] An outline of what I believe to be the most plausible solution is given in Costa 2011, Ch. 5.

[19] These two forms were also noted by Anthony Kenny (1973).

These two forms of transgression have a striking similarity to the psychoanalytic distinction between the two mechanisms of the *primary process* (*primärer Vorgang*), called by Sigmund Freud *displacement* (*Verschiebung*) and *condensation (Verdichtung*). Hence, it is worthwhile to explain this process here very briefly. According to Freud, our thinking can involve two distinct processes: the *secondary process* (*sekundärer Vorgang*) and the *primary process* (*primärer Vorgang*). The secondary process is the typically conscious process of rational thought, in particular, scientific thought. In this process, affective or emotional *charges* (*Besetzungen*) are firmly associated with their respective *representations* (*Vorstellungen*). The primary process, on the other hand, is found in dreams, neurotic symptoms, humor, artistic creation, religion, and… philosophy! In all these cases, emotional charges are not rigidly associated with their respective representations (or thoughts) and can be transferred to different representations, insofar as the latter can easily be associated with the former representations. The primary process is what produces the conscious manifestation of unconscious or pre-conscious thoughts, in the latter case understood as non-repressed and consequently always able to become conscious.

The two fundamental mechanisms of the primary process, *displacement* and *condensation*, are more deeply explained in Freud's *Interpretation of Dreams* (1900, Ch. VI).

Displacement occurs when the emotional charge of a repressed representation is transferred to another representation, which is able to elude censorship and become conscious, thereby releasing its endo-psychic tension into consciousness. We have displacement when representation R1, repressed and therefore unable to become conscious, has its charge transferred to representation R2, able to evade censorship and become conscious. A Freudian example of displacement is the story of a Jewish woman who could not marry the man she loved because he was a Christian. However, she dreamed that she gave him her comb. This is her conscious representation in the dream; but in her unconscious, the repressed representation was the idea of giving herself to him in love. The emotional charge passes from the repressed representation to the non-threatening one, which is able to outwit censorship, becoming conscious as a dream. This makes it possible for the charge to be released into the dreamer's consciousness, bringing relief to the endo-psychic tension.

The mechanism of condensation is somewhat different. Here a representation (or group of associated representations) transfers its affective charges to a partial representation belonging to it, which becomes liberated in consciousness. We can represent this by saying that the charges

belonging to the representations {R1, R2… Rn} are usually condensed in one of them, say, R2, which enters into consciousness, in this way allowing the release of emotional charges into consciousness. One example of condensation would be a case if the woman had dreamed that the man she loves forgot his scarf at her home... The scarf is part of the whole representation of the man, and the emotional charges associated with the whole are condensed in this partial representation and released into consciousness.

It is worth remembering that according to Freud, displacement requires full unconsciousness by being a product of repression, while condensation requires only pre-consciousness (i.e., its representations are potentially but not actually conscious) since it isn't necessarily a product of repression.

Now, an investigation of the two mechanisms by which the internal limits of language are transgressed brings into sharper focus the sometimes noted relation between philosophy as therapy and psychoanalysis (e.g., Wisdom 1969), for it shows that philosophical activity is affected not only by a lack of semantic awareness, but also by unconscious motivations.

Let us see now how the primary process works in cases of confusion arising from linguistic transgressions of normal uses of expressions. By using an expression equivocally, a philosopher shifts the use of this expression, applying it in a state-context of a linguistic practice B, though following the semantic rules that this expression should have in linguistic practice A. This *equivocity* amounts to displacement, since the emotional charges associated with the first use are transferred to a new representation. On the other hand – in what we call *hypostasis* – the philosopher tries to apply an expression that can be used according to the rules of two or more linguistic practices, say, A, B, C, etc. simultaneously in a neutral state-context. It is as if there were a single linguistic practice able to bring together these different uses, adding their emotional charges, when in fact this practice does not exists, and there is no justification to initiate it.

Philosophical examples of these mechanisms can be complicated and difficult to describe, since philosophers, being masters of deception (and self-deception), construct their spider webs of far more abstract and complex material than ordinary mortals can imagine. Hence, I will consider only two very simple examples.

For the case of displacement, consider the following skeptical paradox attributed to the Megarian philosopher Stilpo, denying the possibility of predication. For Stilpo, if I say that Socrates is wise, this is a contradiction, because I am denying that Socrates is Socrates. That is: I can say of something that it is what it is, but if I want to say something more than

this, I fall into a contradiction, for I am denying that it is what it is… The upshot is that all that we can do is to express the identity of a thing with itself or remain silent.

We can explain Stilpo's fallacy as due to a failure to distinguish the 'is' of copula (of predication) from the 'is' of identity. We can distinguish a linguistic practice of type A – in which the verb 'to be' means 'is the same as' (e.g., 'Socrates is Socrates.') – from linguistic practices of type B – in which the verb 'to be' is used as a copula (e.g., 'Socrates is wise.'). However, Stilpo recognizes the verb 'to be' as having only one correct use: that which is found in state-contexts of type A practices. As a result, each time he observes people using the verb 'to be' in state-contexts of practice B, he understands their use as following the rule of use that the verb has in practice A – meaning 'is the same as.' In this way, he equivocally and systematically displaces the real use from practice A to *practice* B. Since he sees that in all these state-contexts of practice B he cannot apply the 'is' of identity typical of practice A, he falsely concludes that true predication is impossible.

I will now offer an easy example of hypostasis in philosophy. Consider this suggestion made by a philosopher, according to whom the verb 'to be' must have a truly primordial sense, which is not only that of copula, but also of identity and of existence together! To justify this, he considered the sentence: 'To be is to be' (*Sein ist Sein*). This sentence says not only that 'to be' has the property of being, but also that 'to be' is the same as 'to be,' and finally that 'to be' has the property of existing (of being).

Against this *folie metaphysique*, a critique of language will tell us that it is much more plausible to think that what the philosopher seeks with the 'is' in the sentence 'To be is to be,' although grammatically correct, is semantically only an incoherent mixture of different uses of the verb 'to be.' These were created for different practical purposes but have no justification when mixed together, except the satisfaction of an *ad hoc* philosophical claim. It is a hypostasis: a condensation arbitrarily mixing three very distinct modes of use – meanings – of the same word in a supposedly neutral state-context. However, these three modes of use belong to three actually distinct practices, say, the identifying practice A, the predicative practice B, and the practice of attributing existence C. In the best case, this is an example of multiple ambiguity; but since the philosopher is claiming to have discovered a way to achieve the primordial sense of Being in a factually arbitrary way, the diagnosis is of mere incoherence and illusion.

I offer these explanations because in criticizing the metaphysics of reference, we very often denounce equivocity and hypostasis. Wittgenstein

suggested that philosophical *maladies* have their origins in a 'craving for generality': in efforts to achieve generalization without sufficient reasons, by reductionist means, usually influenced by the greater success of natural science (1975: 18). We can now suggest that here as well the frequent case of equivocity may also work as a compensatory byproduct of repressing some kind of undesirable awareness.

An additional point is that striving for generalization is inherent in the philosophical endeavor (particularly as revisionary metaphysics) even if it may be ultimately doomed to some kind of failure. Wittgenstein concedes that the philosophically unavoidable bumps up against the walls of language have the mark of profundity (1984c, sec. 111). The reason for this concession is that these confusions, when able to strike us, have the potential to point to relevant issues insofar as they might force us to search for the right way to avoid the illusions they produce in us. As I intend to show, much of the metaphysics of reference is grounded upon the forms of confusion described above, particularly equivocation (displacement), which makes them the right target for the therapeutic critique of language.

12. The form of semantic-cognitive rules

In an approximative way, we can now expose the general form of a cognitive or criterial semantic rule, anticipating what will be considered in more details in the next chapters. This rule is constituted, on one hand, by a relation that can be summarized in the sign '~>,' which means either a strong inductive inference ($p > 0.5$) for empirical knowledge, or a deductive inference ($p = 1$) for logico-conceptual knowledge. By 'C' I mean *the criteria to be satisfied* and, by the result 'A' I mean the (usually non-reflexive) *meaning-awareness regarding the rule's application* and linguistically expressible by a declarative sentence. Here is the basic schema:

$$C \sim> A$$

This schema of a semantic-cognitive rule is too simplified, for the criteria are usually multiple, varied and staggered in complex procedures. The satisfaction of a (definitional) criterion under the state-context of a certain practice should give place to a meaning-awareness, a judgment expressible by an assertoric sentence. The cognitive content or meaning or sense is the whole procedure of rule-following, including still unverified criteria, insofar as they also belong to the same ramified rule.

Complementing what was said, there is a second cognitive element associated with the semantic-cognitive rule, which is the *awareness of the consequences of the satisfied content* – of the applicability of the rule-combination or rule-complex. I believe this could be explained by theories of consciousness such as those briefly summarized at the end of chapter II. It would be, for instance, what has been called the 'availability of content to reasoning and action' (Block), the 'transmission of content for the mind's global workspace' (Baars), 'brain celebrity' (Dennett), etc. It is the full consciousness of what the meaning of a declarative sentence represents.

Calling the meaning or cognitive content [{C1 ∨ C2 ∨… ∨ Cn} ~> A], where each criterion is seen as sufficient for the meaning awareness A, and calling E its cognitive consequences (as the transmission of content to the global workspace of mind), we can summarize a typical common form of a semantic-cognitive rule added to its cognitive effects as follows:

$$\frac{[\{C1 \vee C2 \vee \ldots \vee Cn\} \sim> A]}{\text{Semantic-Cognitive Content}} > E$$

To this, we should add that when the semantic-cognitive rule is the verifiability rule, the cognitive content is the thought-content expressible by a declarative sentence, as the whole summarized in square brackets.

In order to better understand this representation of a criterial rule, suppose that C2 is assumed as the given criterion for the meaning awareness of what can be expressed by the statement 'Calphurnia urged Caesar to stay.' The understanding that Calphurnia urged Caesar to stay at home is a meaning awareness. A regarded application followed by conscious effects E that can be silently thought or spelled out.[20] Now, we can consider two semantically relevant situations:

(a) When we add informative content to the verbal formulation of A, associating it at least potentially with some conventional procedure from which it results, for instance, C1 ~> A. Then we have semantic-cognitive meaning; the rule is regarded as at least potentially applicable. This act of regarding is an imaginary *rehearsal* of the true application. Here the cognitive meaning, for

[20] C. S. Peirce's view, according to which all thought is in signs, seems to be wrong, considering that we are surely able to think without using words. But on second thought, it is plausible that in having these non-linguistic thoughts we are using non-linguistic mental signs, like imagistic and emotivist ones.

instance, e.g., the verifiability rule, is 'put under consideration.' But this does not mean that propositions or thoughts are statically regarded as abstract entities – this does not exist! What is meant is that they are known as applicable or even applied in our imagination, even if in a limited way, only felt as potentially applicable in a real situation, with the result that we make ourselves aware of semantic-cognitive content as a possible occurrence of a rule-in-its-concrete-application. This already makes us to a certain extent aware of the foreseeable effects E, once we are using it as an instrument in a search for possible utilitarian consequences.

(b) When a criterion, such as C2, is contrastively seen as actually satisfied; then we have an application of the semantic-cognitive rule, which can be symbolized as C2 & [{C1 ∨ C2 ∨… ∨Cn} ~> A] ~> E. This fulfilled, A inevitably produces a *true referential awareness*, which should bring about E as A's availability for reasoning and action, its transmission to the mind's global workspace, brain celebrity, etc. given by theories of consciousness, since it is what results from consciousness of a really given factual content. Here we say that the semantic-cognitive rule is *effectively* applied or applicable. In this case, we add to the meaning-awareness A a judicative value, and if we associate this cognitive application of the rule with its spelling, we have an assertion, namely, a statement spelling out a sentence whose content is accepted as true, having as C2 its verifier. Notice that what is judged or asserted is the whole content: the verifiability rule along with the satisfaction of its criteria.

It is interesting to note that there is some proximity between our conclusion and inferentialist approaches to meaning. If we say that a content, a semantic-cognitive rule, is available for reasoning and action, we also mean that the content – which is in itself inferential – would be inferentially open to those related contents. This is what I believe can be understood as the cognitive *effect* of the satisfaction of the semantic-cognitive rule. However, I will not risk mixing this inferential openness proper of the cognitive awareness of content with the real meaning, because this openness is only a consequence of the instantiation of referential or cognitive meaning won through the application of its semantic-cognitive inferential rules.

The usefulness of these sketched formulations will gradually become clear in the course of this book.

13. What is wrong with the private language argument?

I do not believe that there is only one possible interpretation of the so-called private language argument (Wittgenstein 1984c, I, sec. 244-271), a name that isn't even present in Wittgenstein's text. There are a variety of more or less interesting interpretative alternatives. In fact, to interpret Wittgenstein is like trying to assemble a jigsaw puzzle, knowing from the start that some pieces will inevitably be left over.

This isn't a problem for me, insofar as my aim here is not properly interpretative, even if I believe my interpretation is the one most faithful to the central line of Wittgenstein's thought. What I want is to reconstruct Wittgenstein's 'argument' in a way that makes its consequences as philosophically strong as is reasonably possible. This philosophically strong formulation will be important, because if it is right it means the destruction of all our human subjectivity as it is currently understood and as it has been understood in the traditional philosophy (e.g., in the *cogito* or regarding *sense-data*). A private language argument with trivial conclusions would be of scant interest.

I can begin with the contrasting case: public physicalist language. How do we learn to identify and distinguish different types of physical objects? For example: how does a child learn to identify references of the word 'ball'? This doesn't happen by means of verbal definitions, but ostensively: adults point to examples and say things like, 'This is a ball' or 'That isn't a ball'... and the child eventually learns what types of objects are round balls. But this learning is only confirmed when a new ball is presented and the child shows adults that it is able to *re-identify* the object as belonging to the ball type. In this case, based on agreement among other speakers of the language regarding correct re-identification, it is possible for everyone (adults and the child) to know that the child has learned the rule for identifying ball-type objects. That is, the only way to know that we have learned a rule is ultimately to confirm our way of application by interpersonal checking.

Consider now what happens when we try to identify internal mental entities of a phenomenal nature (sensations, emotions). In this case, we cannot do any checking of interpersonal re-identifications. Suppose that a person is expected to learn to identify an internal state, for example, a feeling of pain. Other people cannot teach him to do this, because they cannot know if and when he feels pain or how it feels to him. But let's suppose that independently of any public language a person decides to point inwardly to some feeling and identifies his feeling through a sign that he himself has invented. Suppose this sign is 'P' (for 'pain'). Imagine

now that the next time he feels pain, he says to himself 'P,' intending to point to the same internal mental state. In this case, he won't be able to know if he is really pointing to *the same* phenomenal state that he initially pointed to, because there are no other speakers who can check the correctness of his rule application, that is, who are able to confirm or refute his identification. As Wittgenstein realized:

> 'I impress it on myself' can only mean: this process brings it about that I will remember the connection *correctly* in the future. But in this case, I have no criterion of correctness. One would like to say: whatever seems right to me is right. And that only means that here we cannot talk about 'right.' (1984 sec. 258)

Where interpersonal correctness criteria cannot be found, we cannot distinguish between following a rule and the mere *impression* of following a rule. However, this distinction is indispensable, because without it we have no way to construct something that we may effectively call 'a rule.'

Since language is a system of rules, the generalization of this result leads us to the radical conclusion that there cannot be a language whose objects of reference are internal phenomenal states. For Wittgenstein the only construable psychological language seems to be the one based on behavioral *expressions* of internal states, transforming expressions like 'Ouch!' into 'I feel pain.' (1984c, sec. 244) Nevertheless, he concedes the existence of these mental states, rejecting behaviorism. This is in my view a clearly incoherent move, since under his assumptions real mental states should be beyond the reach of linguistic rules and therefore cognitively inaccessible, not expressible in language and in the end senseless… once he also writes that something about which nothing can be said has as much value as nothing (1984c, sec. 304).

The problem, as Ernst Tugendhat once noted, is that the private language argument is too counter-intuitive to be correct. The point, however, is to discover where the argument's weakness lies. In order to find this, we need to make two things clear. The first is that we will probably only stop regarding a rule as a rule if we conclude that it is *logically* impossible to be corrected. A rule does not cease to be a rule just because for some contingent reason it was not in fact interpersonally checked. After all, it is beyond doubt that many of the rules we follow, for one reason or another, have never been interpersonally checked. I can invent for myself the rule of never eating creamed spinach, and nobody needs to be informed of this rule. There are rules that for merely circumstantial reasons cannot be checked, such as those made by a

shipwrecked sailor who is never rescued and consequently lives and eventually dies alone on a remote uninhabited island.

An objection that could be made to this interpretation is this: Wittgenstein's argument demands that any rule, in order to be a rule, must be publicly checked for correctness, and not just be *able* to be publicly corrected (correctable). Even if this interpretation were true, it would be utterly uninteresting. For it expresses only an extremely implausible and methodologically anti-Wittgensteinian idea, jeopardizing our common sense certainty that there are too many rules that we follow that have not been checked by others. In fact, if we wish to overstate skepticism, we could also argue that no rule can be applied in situations where it cannot be subjected to *simultaneous* interpersonal correction – after all, there is no guarantee that in the absence of this control the rule will be correctly interpreted and applied... However, gratuitous forms of skepticism like these are too implausible to persuade anyone.

With this in mind, let us now interpret Wittgenstein's argument as assuming that the rules of a phenomenal language must be *logically incorrigible*. Let's suppose that every morning when waking up I unintentionally follow the rule to remind myself of the first sentence of Dante's *Divine Comedy*, but that I always immediately forget what I have done. Here we are already close to nonsense, and we would reach total nonsense if it could be proved to be logically impossible to know if this happens.

We conclude that it is the assumed *logical incorrigibility* of phenomenal language that definitely convinces us of the plausibility of the private language argument: it seems very plausible to assume that a rule that logically cannot be corrected cannot be considered a rule. If the rules of our (supposedly) private phenomenal language are logically incorrigible, it seems that they cannot, ultimately, be distinguished from mere impressions of rules.

This reconstruction of Wittgenstein's argument is not only the most interesting and reasonable. It also uncovers what I believe to be an important implicit assumption made by him. Once he noted, for instance, that even though person A's nervous system could be connected to that of person B, so that A could feel a wasp stinging B's hand, only the location of pain would be shared, but not the pain itself. This is because pain felt by A would be A's pain, while pain felt by B would still be B's pain (Wittgenstein 1975: 54). In his most famous article, surely read by Wittgenstein, Frege noted that if another person could enter our minds to observe a visual representation, the representation he experienced would be *his own* and not ours (Frege 1892: 30). Now, this kind of consideration

leads to a dogma generally assumed by earlier Twentieth Century analytical philosophers: the thesis that *phenomenal states are logically non-shareable*.[21] If this thesis is correct, then interpersonal corrigibility of phenomenal language would be logically impossible, which seems to be a reasonable ultimate foundation for the private language argument.

At this point is understood, all we need, if we wish to destroy the private language argument's ultimate foundation is to show that the logical non-shareability of phenomenal states is a false principle! That is, we need to show that although the rules of a phenomenal language have never been interpersonally corrected, they *are* – contrary to what Wittgenstein and many philosophers assumed – *logically corrigible from an interpersonal perspective*, this being the hidden flaw that tacitly supports the private language argument.

It's hard to imagine a thought-experiment able to prove that phenomenal states are logically shareable. We can begin by making an analogy with computers. Suppose A and B are updated versions of the primitive kind of automata called by Grey Walter *machina speculatrix*, which fed on light and spent all their time in search of it. Suppose automaton A meets automaton B, and that A is able to *read* the information content that B has accumulated in its searching. Although automaton A can copy these data first, and only afterward read them in its own system, so that such 'contents of experience' become an unshared part of itself, there is no contradiction in thinking that A can read these 'contents' *directly* in B, as if they were its own, thereby sharing them with automaton B before selecting relevant data! This would, in fact, be the simplest and most direct method. Why should we think that in an analogous situation we humans would need to be different from machines?

Perhaps it is even possible to imagine that someday there will be two human beings, A* and B*, who somehow are able to share some functioning of parts of their brains. Suppose that their limbic system is in some way tuned essentially the same, while the cortical regions of A* and B* remain distinct. Now, it seems conceivable that a mental state of pain that occurs in relevant parts of this one and same limbic system could be shared by subjects A* and B*, even though their conscious interpretation of pain, made in their distinct cortical regions, are qualitatively different. If we understand pain as essentially a process occurring in a limbic system,

[21] See, for instance, A. J. Ayer 1972: 196.

then A* and B* really could share the same pain, demonstrating possible interpersonal checking of the same internal phenomenal state.[22]

The thought-experiments considered above suggest that it is logically possible to distinguish:

(a) the *subjective interpretation* of a phenomenal mental state *X* from
(b) the phenomenal mental state *X in itself*.

In fact, this separation seems possible. We know cases of hypnosis where people are led to feel pain even though a source of this pain is absent or not to feel a real pain. We know the case of a patient at the dentist who, because he is afraid of treatment, believes he feels pain when he really only feels the sensation of friction…

Now, if we accept that it is logically possible to separate (a) and (b), then the interpersonal sharing of mental phenomenal states turns out to be logically and maybe physically and practically possible as well, which at least in principle warrants the possibility of the interpersonal checks of identification rules for mental states. In this case, the private language argument fails because the logical non-shareability of phenomenal states is a false principle. In this case, the rules of phenomenal language acquire an epistemic status that does not essentially differ from that of the rule I made for myself of never eating creamed spinach; in principle, both rules could be checked. Consequently, we are entitled to assume that what we now *believe* to be the rules of our phenomenal language may, in fact, be the actual rules, since they are at least logically susceptible to interpersonal correction.

Furthermore, we are also entitled to say that the correction of rules for the identification of phenomenal states is *highly probable*, since this probability is very well confirmed in an indirect way by a multitude of systematically related associations between interpersonally accessible physical phenomena and reports of internal phenomenal occurrences. For example: if wrinkling the forehead is often associated with the statement 'I

[22] In fact, I think we are not very far from this result than some might believe. Computational fMRI brain reading is already close to being able to reconstruct mental states (images, intentions, memories), making them interpersonally graspable as well for the person who is having these states: you can see your own mental images (visual sense-data) represented on a screen, and others can see your represented images on the same screen. (e.g., Nishimoto 2011) Even if they are not the images (visual sense-data) in themselves, the experiment already suggests that your consciousness of these images is detachable from them.

feel pain' when one believes one has a feeling *x*, wrinkling the forehead indirectly reinforces the probability that when applied the words will really refer to the same feeling – even if only minimally. Against this kind of reasoning, Wittgenstein imagines a situation in which when he believes he has the (non-perceptible) particular sensation P the manometer always shows that his blood pressure has increased. This assures a correlation between his subjective thinking that he has the feeling P and an increase in his blood pressure; but it does not guarantee that the rise in his blood pressure will be correlated with *the same* sensation on various different occasions (1984c, sec. 270). Indeed, it does not guarantee that the last correlation will be the same, but we feel that in a small measure it *increases* the probability that P is being correctly correlated with the same feeling. As we normally have a very great interweaving of such correlations, what we normally make is a well-grounded reasoning by analogy, allowing us in the end to reach a very high probability of associating something like the belief that we are having what we call 'pain' with the pain that we really feel. The difference is that in normal cases of reasoning by analogy we can do a final check to prove that the inference was correct, while in the case of subjective, inner feelings this seems impossible. But if we can do this in principle – if the principle of the logical unsharability of mental phenomena is not true – there is no justification to question our reasoning by analogy regarding our feelings.

It seems clear that our reference to internal phenomenal states is not essentially different from the case of the conclusion based on a large amount of convincing circumstantial (indirect) evidence, that a certain woman was in fact murdered by Jack the Ripper, even though the true identity of this serial killer was never and might never be proved. Even if no one actually saw the woman being murdered, the details of the murder and all the circumstantial evidence that taken together point to this very peculiar murderer are already highly convincing.[23]

14. Concluding remarks

Returning to our initial question about the nature of the intermediate link, we can now see more clearly why and how the intermediate link between words and things can be read in two different complementary modes. These are the psychological mode, which considers some particular cognitive bearer of the link, and the semantic mode, in which particular bearers of a link with their psycho-physical particularities are left aside.

[23] Costa 1997, 433-448; *Cf.* also Costa 2011, Ch. 5.

That is: cognitive meanings are semantic-cognitive rules that can be considered in their possible or effective application and that when regarded from the viewpoint of their conditions of satisfaction, can be seen as semantic-cognitive criterial rules. As will be made plausible in Chapter V, the cognitive meaning of a statement should be nothing but a *verifiability rule* that really applies when some criterial configuration required by it is adequately satisfied, making the statement true; it being otherwise false.[24]

Nonetheless, it is important to maintain a clear distinction between the semantic and the psychological aspects of the intermediate link, as philosophers like Frege and Husserl insisted, even if they did it in a needlessly equivocal way. The semantic aspect is conventionally grounded and grammatically necessary; the psychological aspect is spatiotemporally given and in its psycho-physical particularities contingent. But contrary to what these philosophers have supposed, nothing semantic can really exist *outside* of cognitive instantiations. Semantic entities are nothing more than conventional structures that exist only when embodied in mental acts, in applications of rules, even if considered in abstraction from their contingent bearers. To assume that semantic entities can exist without any psychological basis is to hypostasize their nature.[25]

[24] Note that there are non-referential rules: we can not only have rules that relate (a) the empirical data to cognitions, but also (b) cognitions to other cognitions, and (c) cognitions to actions. But concerning the issue of reference, what matters is the first kind of rule, which is responsible for cognitive/referential meaning.

[25] As I see it, there is a great variety of ways to make this hypostasis. One of them is to identify sense/meaning with Platonic entities (Frege, Husserl); another (already criticized in the Appendix to Chapter II) is to identify meaning with something external like essences of things (Putnam); another is to identify meaning with minimum units of reference (Russell); and yet another is the attempt to identify meaning with psychological communicative intentions (Grice).

APPENDIX TO CHAPTER III

TROPE THEORY AND THE UNSUSTAINABLE LIGHTNESS OF BEING

> Any possible world and, of course, this one, is completely constituted by its tropes.
> —*D. C. Williams*

'Could you show me some *properties* (qualities, characteristics…) of the things around us?' Asked in this way, any normal person would surely point to a few nearby objects, naming their properties (qualities, characteristics…), such as the redness of this sofa, the hardness of that wall, the property of a shirt of being made of cotton… Many traditional philosophers, however, would say that these things cannot really be properties in the true sense of the word. For in this true sense, properties are abstract entities, universals accessible only to our intellect, not to our senses.

This comparison suggests that the ontological starting point of traditional realism, particularly in the form of Platonism is opposed to the ontological starting point of ordinary people and even of our own modest common sense. Common sense begins by considering as prototypical examples of properties the spatiotemporal properties directly given to us in perceptual experience, only afterward considering those properties that are in some way derived from perceptual experience. The contemporary ontology that shares this commonsense view is called *trope theory*. Properties are for trope theorists spatiotemporally located entities called 'concretized properties,' 'particularized qualities,' 'individual accidents,' 'quality-bytes,' 'abstract particulars' or simply 'tropes.' According to trope theory, universal properties should follow from the ontological building blocks that are the spatiotemporally particularized properties or p-properties called tropes, and not the other way round.

One reason for the importance of trope theory resides in the fact that since the rise of nominalism already in the Middle Ages, this might turn out to be the only really groundbreaking advance in ontology. Although the concept of trope as a particularized property has been known at least

since Aristotle, only in 1953 did an American philosopher named D. C. Williams conceive of the bold idea to assign tropes metaphysical place of pride as the universe's only fundamental ontological building-blocks.[1] His central aims were to use the notion of tropes to solve (or dissolve) the traditional problem of universals and to explain the nature of concrete particulars. In fact, pure trope-theory is a one-category ontology. Because of this, my hunch is that the theory of tropes is so revolutionarily simple in its fundamentals that it could produce an upheaval in ontology similar to that caused by the introduction of new physicalist theories to solve the mind-body problem in the second half of the twentieth century.

In what follows, instead of doing the hard work of discussing different versions of trope theory, I will take the easier and more direct route of outlining the view that from my assumed methodological perspective seems more plausible, namely, a methodology that gives primacy to established knowledge (Ch. II, sec. 5).

1. Introducing Tropes

First, what are tropes? Although tropes (or properties) considered as simple cannot be intrinsically defined, they can in my view be clearly characterized as follows:

> Tropes (*Df*): are *properties localizable in space and enduring in time, regardless of their vagueness*.

As such, these particularizing properties can be identified as the empirical *designata* of predicative expressions. The most obvious tropes – fundamental from a genetic-epistemological perspective – are those accessed by direct perceptual experience, like *qualities*. Examples of quality-tropes are the yellowness of this sofa, the heat of that stove, the smell of a particular daisy at a certain time and the song sung by a particular blue whale to attract a female. Other tropes would be the red color of the Golden Gate Bridge, its weight, hardness, form, height above sea level… These are all that we could call *external* (third-personally

[1] This groundbreaking work was D. C. Williams' paper 'The Elements of Being' (1953), because he was the first to propose constructing the whole world using *only* tropes as elementary building blocks. The most relevant attempt at a systematic development of trope theory remains in my view Keith Campbell's book, *Abstract Particulars* (1990). Since then, the discussion devoted to this view has grown steadily. For access to the literature, see Anna-Sofia Maurin's, 2013.

accessible) physical tropes. However, tropes can also be *internal*; they can be psychological properties, like a feeling of pain, sorrow, love, and pleasure and even a whole mind, insofar as not understood as a thinking substance (Williams 1953 I: 17). They can be partly internal and partly external like a belief, emotion, purpose, love affair, act of contrition or expression of impudence (called by Williams mixed tropes); and they can be events like a smile, sneeze, election, cold snap, triangle, circle, shape or bodily form (Williams 1953 II: 171 f.). We can prove the reality of tropes by considering that they can be removed, like the color of a cloth (Campbell 1998: 352) and can be objects of selective attention (Loux 2002: 86): gazing at the ocean, one can alternately concentrate on its color-tropes, the form-tropes of its waves or their sound-tropes.

Simple tropes appear in combination with other tropes, and some conglomerates of different kinds of tropes are highly complex and multifarious. This is the case of biological properties like that of a certain cat being a mammal. This is the case with some psychological properties like Céline's idiosyncratic personality. And this is also the case with social properties like that of India being a democratic country. If I say that India is a democracy, 'being a democracy' is a property-trope dependent on the country, the individual entity called 'India,' though this trope is surely a very complex one. And there are complex and diversified cultural properties like the socio-cultural traits emphasized in the ancient Spartan state. In all these cases, the tropes are in various ways spatiotemporally located, and they are properly referred to with predicative expressions (they are at least logically repeatable).

Tropes contrast with what I prefer to call *individuals* (objects): these are things that are seen as unique and non-repeatable and are referred to by singular terms like 'this daisy,' 'that blue whale,' 'the Golden Gate Bridge,' 'Socrates' and 'India.' In the standard case they are what can be called 'material objects' and, as we will see, nothing but *compositions* of tropes. However, some compositions of tropes are individuals without being material objects. This is the case of a rainbow or of a cloud in the sky. And there are individuals that are constituted by the absence of tropes, for instance, a particular shadow.

Moreover, there are complex tropes like a performance of Beethoven's Fifth Symphony, which are homogeneous in the sense that they consist of only one kind of trope comprising a great diversity of sound-tropes. They can be designated by means of a predicative expression, as in the statement 'The orchestra performed the Fifth Symphony.' Considering that the Fifth Symphony can be performed over and over by many different orchestras at different times and in different places, it is clear that it is

better classified as a repeatable complex homogeneous trope and not as an additional individual; moreover, it is dependent on an orchestra (an individual) to be performed, while individuals (e.g., the Vienna Philharmonic Orchestra) are relatively independent in their uniqueness.

Finally, one can consider the existence of *indirectly accessible*, derivative tropes. This would be the case of fundamental physical forces: in order to have a clue about them, we need to begin by experiencing our more modest perceptible quality tropes. The fact that these forces are indirectly accessible is only a contingent one (some birds navigate using the earth's magnetic field). This is how things are, even if from the perspective of physical science the origins could be reversed.

As particularized properties, tropes have identity conditions. As an attempt to clarify this, I propose an ontological condition (a) followed by a linguistic indicator (b):

Tropes are identified:

(a) By their *spatiotemporal existence* to the extent that they display *sufficient continuity* over space and time and are amenable to certain direct or (often) indirect experiential ways and conditions of access, and

(b) By being linguistically designated by *predicative expressions* of singular statements whose nominal terms refer to individuals.

So understood, tropes contrast mainly with individuals such as material objects referred to by means of nominative expressions, particularly proper names.

The linguistic indicator (b) has a guiding function: as spatiotemporally located properties most properly linked with individuals, tropes are usually designated by means of predicative expressions like '…is red.' This isn't always so straightforward: in statements beginning with demonstratives like 'This is a daisy' or 'There is the Matterhorn' it is preferable to take the nominal terms (indexicals) 'this' and 'there' as referring to spatiotemporal places, as localizing rules for the identification of the individuals daisy and mountain, which justifies the non-application of the linguistic requirement (b) to its supposed predicates, whose owners are in fact individuals and not tropes. Better to analyze these sentences relationally as '<This place is> where <a daisy> is located,' and '<That is the place> where <the Matterhorn> is located,' sentences in which the predicate designates the property-trope '*x* where *y* is located.'

Regarding the ontological condition (a), I have something more to say. Consider the following example: the pair of shoes I am wearing is brown.

The right shoe's property of being brown can be seen as a trope, since it displays continuity and is located on my right shoe, and the left shoe's property of being brown can be seen as another trope since it displays continuity and is located on my left shoe. Because these shoes have different spatial locations, we can regard them as displaying two tropes of the color brown. And because of the relatively homogeneous continuity of the right shoe's color, this color can be said to be only one trope – a (located) property. The smoothness of my left shoe is also a trope that has the same location, homogeneity and maybe even the same duration as its brown color. Does this mean that this brown and this smooth are the same trope? No, since they are accessed through different forms of perception and under different conditions. This is the most natural way to identify properties, although there is much more to be considered on this point.

To the further question of how much my left shoe's trope of brown can be subdivided, one possible answer would be: into as many unities as we can distinguish. However, since depending on perceptual distance and acuity we can distinguish different amounts, this does not seem to be very helpful (*Cf.* Campbell 1990: 136-7). Because of this, and again drawing on common sense and natural language, it seems better to say that the *unity* of a trope – which we can rightly call a property – is usually better established by the natural limits of its spatiotemporal continuity and what is considered as being the same, disregarding its possible divisions. Thus, for instance, the whiteness of a wall would be a myriad of tropes if any visible point of whiteness were considered a trope; but considering a trope of whiteness to be a continuous whole, we are not only being economical but also following the usual linguistic practice. Indeed, we would rather say that this wall 'has the property of being white' than that it has a myriad of punctiform properties of whiteness. The size and form of the wall, on the other hand, also deserve to be called tropes, since they can be spatiotemporally located. A related question concerns the duration of tropes. How long will my left shoe's brown trope last? A reasonable answer is: it will probably survive no longer than my left shoe. A trope lasts as long as it remains essentially the same, maintaining its spatial continuity.

I mention all these seemingly trivial things because hasty considerations can easily give rise to attempts to discredit identity conditions for tropes, for example, by pushing precision beyond its contextually reasonable bounds. The vagueness of our identity conditions for tropes is as much a direct consequence of the way we experience the world as of the way the world is supposed to be under our assumed practices, enabling us to define a conceptual system with a suitable degree

of precision. Moreover, many complex tropes (e.g., socio-historical tropes) can be highly dispersed in space and time. This makes their boundaries still less determinate.

Since tropes are any spatiotemporally situated properties, they are also existent particulars. This is because existence – as we will see later in this book – can be seen as the effective applicability of a predicative ascription rule to at least one thing. By asserting existence we assume a need to spatiotemporally locate a trope or a set of tropes. Moreover, tropes are said to have proper existence, though I must disagree with Keith Campbell's view that their existence is independent (1998: 353). He gives as examples the blue of the sky and the colors of the rainbow. However, the blue sky above must be identified against the landscape below, and the colors of the rainbow are intransitively related one another and form an arc against a certain background, and all these things are, according with our definition, tropes. Therefore, I would prefer to say that tropes have rather an *interdependent* existence.

Are *spatial forms* and *duration in time* tropes? Well, these things cannot be found without being associated with tropes, a shape with a color, a volume with a weight, a duration in time with the continued existence of some tropes or clusters of tropes... Campbell, disagreeing with Williams, did not consider forms as tropes because of their dependence upon other tropes (Campbell 1998: 360-361).[2] However, as I noted above, his examples are inadequate: tropes have to be always to some extent interdependently considered. If we hold this view together with our definition of a trope as any spatiotemporally localizable property, we can see forms and durations as limitations in space and time respectively. They would arise from limitations imposed by standard quality-tropes. Hence, it seems that we could view forms and durations as kinds of tropes. Let us call them *limiting tropes*.

Another question is whether *relations* are tropes. Since relations are spatiotemporally located, though often only in a rather vague way, and since relations are designated by means of dyadic or polyadic predicative expressions, it seems clear that relations are tropes, even if their existence is subsidiary to the existence of their *relata*. Although there are different kinds of relations with different strengths, particularly important is the causal relation. For instance: 'The throwing of a stone broke the window.'

[2] In his book on tropes, Campbell writes, 'because boundaries in space need to be drawn rather than revealed it is perhaps best to view individual specimens of each of the shapes as quasi-tropes rather than as genuine tropes.' (1990: 91) This argument is not forceful since a conventionally charged intromission of epistemic subjects is inevitable in any conceptual application.

As Williams and Campbell have noted, a causal relation should be analyzed as a relation between tropes (Campbell 1990, Ch. 5.15). The relational predicate '*x* causes *y*' is not between the objects stone and window but between a cause, such as throwing (a stone), and an effect, such as breaking (a window). Cause and effect are here located *events* associated with different individuals, which can be represented by means of statements ('The stone was thrown', followed by 'The window was broken'), being all made of tropes according to our identity conditions. It is doubtful if a causal relation is internal. We define an internal relation as a relation that exists as a consequence of the existence of their *relata*, so that if the relation does not exist the *relata* will be different. But a trope-event *x* will only be a cause of *y* if the right contextual conditions are added, what must be extrinsic to the *relata*. A straightforward case of an internal relation, however, is that of strict similarity between two tropes, which I understand as a relation of qualitative identity. For instance, 'The blue of this ocean is like the blue of the sky above it.' Once these two blues are given, the similarity follows. Moreover, it may not be as easy to admit, but the relation of strict similarity is also not just predicatively designated; it is also spatiotemporally located: it is in-between and not out there. Therefore, it should also be classified as a relational trope, even if subsidiary to its *relata*. Like causality, strict similarity is in this way a *dependent trope*.

One objection to the idea that relations are tropes could be that if relations are tropes then the relational trope and its *relata* must be related by a new relational trope, and so on *ad infinitum*. We can argue against this objection by first noticing that the same problem comes up again in a stronger form in the case of one-place predications. In other words, if *a* refers to an individual and *b* refers to another individual, and there is a relation *aRb* so that this relation produces an infinite regress, then the same should be true of a one-place predication of the form *Fa*, as in the statement 'The Earth is round.' That is, we would need a relation **R** to relate the object referred to by the nominal term 'the Earth' and the trope of roundness designated by the predicate '…is round,' symbolizing it as F**R**a. Being related to the *relata* F and a, this relation **R** would require two new relations 'FR1**R**R2a', and so on *ad infinitum*. But this seems preposterous! The strangeness becomes clearer when we replace the symbols with words and see that we fail to give a sense to these new relations. It does not make sense to say 'The Earth is related to its roundness,' instead of saying 'The Earth is round.' Hence, it is more reasonable to see the link between subject and predicate as what some philosophers called a 'non-relational tie' (Strawson 1959, part II, Searle

1969: 113), something like the invisible link of a chain, to use Wittgenstein's metaphor. They are not tropes but pseudo-additions in a literal sense of the word. Thus, we do not need to postulate *F**R**a* in order to explain *Fa*.[3] And if this seems obviously true of the monadic links represented by singular predicative sentences, there is no reason not to extend this result to the relations said to produce a regress. After all, relations must be seen as linked with their *relata* in the same way as non-relational properties are linked with their objects. To see this clearly, consider the following example: (i) 'Socrates is a friend of Plato.' Since friendship is a relation, one would be entitled to replace sentence (i) with (ii): 'Socrates has a relation of friendship with Plato,' which still says the same thing by being interpreted as specifying that the kind of relation is that of friendship. But if we try to go ahead, deriving from (ii) the sentence (iii) 'Socrates relates himself to his relation of friendship, which is itself related to Plato,' which is an instantiation of *aR1**R**R2b*, we again wind up speaking nonsense.

2. Tropes and Universals

The theory of tropes is important because it promises a parsimonious solution for at least two perennial ontological problems: the problem of universals and the problem of concrete individuals.

I begin with the problem of universals. Linguistically stated, this problem consists in the question of how we can apply a single general term to many different individuals; ontologically stated, it consists in the question of how it is possible that many different individuals can share *the same* property. Traditional realist philosophers supposed that the only possible solution to this problem is to postulate that a general term refers to a universal understood as an *abstract entity* (existing *ante rem* or even *in rebus*, according to the 'Platonist' or the semi-Platonist 'Aristotelian' versions of realism respectively) that in some obscure way can be instantiated in many individuals.

For the Platonic realist, we can think and see that this rose and that strawberry are red because they instantiate or exemplify the idea (universal) of redness ('red-in-itself'). For Plato, the world was real only insofar as it instantiates ideas. However, this view was never satisfactorily rescued

[3] In Russian, there is no proper verb for the copula. One uses expressions like 'Me nice', 'You beautiful'… Thus, it seems that Russian speakers are less susceptible to such worries.

from unsolvable problems.[4] After all, universal properties must be non-empirical abstract objects accessible only to the intellect. This duplicates the world: we have our empirical world and a world with an infinite number of abstract entities whose intelligibility is highly questionable and for which we have no identity criteria. Moreover, the realist is left with unsolvable problems of how to explain the supposedly causal relation between these abstract entities and our minds. Finally, as we already noted, if you ask a layman where properties *are*, he will answer by pointing to the blue of the sky, the hardness of a table, the softness of jelly… and not to an otherworldly Platonic realm.

This contrast leads us to the suspicion that only a disposition originating from the pressure of some mystical or quasi-mystical belief could lead to a committed Platonic solution. It exemplifies the consolation of what a Nietzschean philosopher would call a 'world of beyond' (*Überwelt*). Philosophers are particularly susceptible to this sort of thinking; they are to some extent unworldly creatures, and it may be a temptation to adjust their minds to see properties in such an idealized way.

The Aristotelian solution was an attempt to bring the Platonic archetypal ideas down from their heaven (the *topos hyperuranion*) to the concreteness of the earth. However, this seems an incoherent middle way. For him universals exist in the visible world so that if there were no world there would be no universals. Now it seems completely impossible to understand how the universal can preserve its unity if its only reality consists in being multiply instantiated by entities belonging to the real world.[5]

Dialectically opposed to realism was nominalism. According to the philosopher Roscelin (XI century), called the originator of nominalism, a universal is a mere *flatus vocis* (emission of a sound), since a general term has no *designatum*. This and similar counter-intuitive views were justly nicknamed 'ostrich nominalism.' A more sophisticated form is the contemporary set-nominalism: a predicative expression designates the set

---

[4] Plato was the first to see some main difficulties of the doctrine in the first part of his dialogue *Parmenides*. Others were added by Aristotle in the *Metaphysics* (book VII) and by later critics.

[5] Although traditionally labeled 'Aristotelian', this is the most simplistic interpretation. More sophisticated interpretations tend to see Aristotle as identifying his forms (ideas) as 'this so-and-so,' the species building the substantial form or essence of the individual (to be distinguished from its matter). According to medieval interpreters, such a form cannot really be a universal; consequently, it is a work of the intellect to abstract the universal from the particular, so that it exists only *post rem*. (Copleston 1993, vol. I: 306; see also Shields 2007, Ch. 6.6)

of individuals to which it applies. This is less counter-intuitive than strange. One problem with this view is that predicative expressions with the same extension – like '…animals with kidneys' and '…animals with hearts' – must mean the same thing since they form the same set. One alternative is to suggest that a predicative expression designates the sets of individuals to which the predicative expression applies in *all* possible worlds (Lewis, 2001: 51). This liberates us from the objection of identities of extensions of different general terms because there are possible worlds where some animals with kidneys have no hearts and vice versa… However, it also leads to implausibility, like accepting the reality of merely possible worlds and assuming the existence of unicorns.

As the solution to the problem of universals by means of realism is too obscure and by means of nominalism is too implausible, trope-theory appears to be the safest lifeboat. To solve the problem of universals by appealing to tropes, we need to introduce the idea of *similarity*, or *resemblance* or *likeness* between tropes, which possibly could be understood as a kind of relational trope. Philosophers like D. C. Williams (1953 I: 9) and Keith Campbell (1998: 358) saw universals as *classes or sets of precisely similar tropes*.

Thus, the universal 'red' refers to the set of all tropes of red, which are unified by the fact that these tropes all have the internal relation of being precisely similar one with the other. For Williams, when we say, 'This rose is red,' we mean that this rose has a red trope that belongs to the set of red tropes; and when we say 'Red is a color,' we mean that the set of all tropes of red (universal-R) is included in the set of all tropes of color (universal-C).

However, there are problems with this view. First, there is a problem with the notion of set or class; if we see a set as an abstract object, it seems that we are abandoning the great advantage of trope theory. Second, there is a problem with size: a set can become larger or smaller; but a universal cannot change its size, for it has no size. It does not help to appeal to an open set, since even open sets also have their sizes, though unknown and also variable… Third, we can develop objections of regress concerning precise similarities based on Russell's criticism of Berkeley's and Hume's nominalism. According to Russell, two patches of the same color have a *relation* of *color-likeness* that seems to be a universal or abstract idea… It is true that a nominalist can decide to consider applying the same analysis to color-likeness, considering it a particular. But then he will face the following problem:

> We may take a standard particular case of colour-likeness, and say that anything else is to be called a colour-likeness if it is exactly like our

> standard case. It is obvious, however, that such a process leads to an endless regress: we explain the likeness of two terms as consisting in the likeness which their likeness bears to the likeness of two other terms, and such a regress is plainly vicious. (Russell 1994: 111-112)

To offer a more detailed explanation, I begin by assuming that likenesses or *strict similarities* are also tropes, as I have assumed before. It must be a case of what I prefer to call 'strict similarity,' because mere similarity or resemblance or likeness lacks transitivity: If trope T_1 is only similar to trope T_2, and T_2 is only similar to T_3, then it is possible that T_3 is not similar to T_1. The solution is to appeal to strict similarity understood as the same as *qualitative identity*, which is the case of an identity between differently spatiotemporally located things (differing from *numerical identity*, which is the identity of a thing with itself). Qualitative identity does not need to be perfect: our cars are both yellow, but your car's color is faded. We must, however, establish a corrigible limit to the differences. Corrigible differences are usually found within the range of a concept's applicability (e.g., turquoise blue and cobalt blue are both called blue) insofar as we have a correction criterion (in the case of blue it is what we identify as corresponding to wavelengths between 450 and 495 nanometers).

Now, according to the kind of reasoning adopted by Russell, in order to construct the set of strictly similar tropes, we need to know that a first trope of identity is like a second trope of identity. But how do we know this? Well, since it cannot be known by appealing to the abstract idea of identity, it must be by appealing to another trope of qualitative or strict similarity. Since the same question can be posed regarding the strict similarities between these strictly similar tropes, it seems clear that this leads to a kind of pyramidal infinite regress.

Russell would see this regress as plainly vicious. Even if this is not the case, I see this as a pseudo-problem born from the wrong solution. And the reason why I think so is that this seems not to be the real way in which we conceive universality. In fact, we can overcome Russell's objection in a much easier way, simply by dispensing with his fixation on classes. The much better way I propose to build universals only from particulars is inspired by just the kind of treatment that particularist philosophers like Berkeley and Hume gave to ideas or impressions in order to ensure their unity. In its plain form, the insight is clearly expressed by George Berkeley in the following passage:

> ...an idea, that if considered in itself is private, becomes general by being made to represent or be in the place of all other particular ideas of the same

> type. ... a private line becomes general by being made a sign, so that the name *line*, which considered absolutely is private, to be a sign is made general.' (1710, Introduction, sec. 12)[6]

Following a similar line of thought, we can symbolize as T* any trope that we wish to use as a pattern or model. Then we can define the universal in a disjunctive way as:

> Universal (*Df.*) = A given trope T∗ *or*… any further trope T that is strictly similar to T∗.

To explain this definition better, we must note that used as a model trope, T∗ in no way needs to remain always the same trope. On the contrary, one can choose any trope T strictly similar to a chosen T* and then use it as a new T∗ in order to make new comparisons. Each speaker is free to use his own T∗ as a model to build the universal. Moreover, what we normally know of T∗ in real life is only some recollection in our memory.[7]

Accepting this definition, we do not need to appeal to sets or classes of strictly similar tropes or some mereological sum to explain universality since the *definiens* covers any trope strictly similar to T∗. The problem of size disappears, since how many tropes are qualitatively identical to T∗ is a matter of indifference. When a person utters the sentence 'This rose is red,' he means that this rose has a trope of red Tr_1 that is identical to some trope of red Tr∗ taken as a pattern (recalled in the person's memory) or any other strictly similar trope. When he utters the sentence, 'Red is a color,' he means that any trope strictly similar to Tr∗ is also a Tc∗ or any other trope strictly similar to Tc∗, as the wider pattern of the color trope. Finally, Russell's problem also disappears, since we don't need to compare one identity trope with another, but only the tropes T_1, T_2,… T_n individually with some chosen trope T∗. Instead of possibly generating an infinite pyramidal regress, the sequence of our comparisions will take the form $T_1 = T*$, $T_2 = T*$… $T_n = T*$, without any need to consider the totality

---

[6] See also the more sophisticated but also less clear view of David Hume (1738, Book I part 1, sec. VII).

[7] We can imagine circumstances in which people are unable to retain memories of the color-trope, but bring with them templates with patterns T* of this color-trope, so that they can compare these patterns with any trope they come across. Moreover, the templates can have the most varied shades of a single color, say, blue. They may call the possibilities that might result from their comparisons 'the universal of a blue color-trope.'

of T's. In other words, as long as all we need to do to get a universal is the ability to compare any given trope with our chosen model trope T∗, there is no need to compare similarities *with* similarities, thereby generating further similarities *of* similarities. Russell's problem does not arise because our particularist definition makes universals mere potentialities instead of actualities.

Furthermore, we can also construct the *universal* 'strict similarity' requiring that some chosen trope Ts∗ (a model trope of strict similarity) is taken as a standard and allowing it to be compared with any other trope of strict similarity strictly similar to Ts∗. Our sequency of comparisions would be Ts_1 = Ts∗, Ts_2 = Ts∗… Ts_n = Ts∗, where Ts∗ can remain the same while other tropes of strict similarity are changing. This means that we have second-order strict similarity tropes referred to by the third-order strict similarity signs '=' occurring between Ts_1 and Ts∗, between Ts_2 and Ts∗, and so on – call them Tss_1, Tss_2, etc. Thus, in order to make reference to the universal composed of these *strict similarities of strict similarities*, we need to appeal to a standard trope of strict similarity of strict similarity Tss∗, and it is easy to predict that we could in principle refer to an indefinite number of higher-order strict similarity tropes by taking this ascending path.

Would this be a vicious regress? I don't think so. For nothing prevents us from *stopping where we wish*, *insofar as we see no reason for going further* – a point that can be understood in terms of explanatory demand. If we do not see any explanatory advantage in going further, we can simply stop where we choose, which is not possible with vicious infinite regresses. A similar consequence results from Platonic realism. As H. H. Price noted (1953, Ch. 1): *the idea of ideas* constantly used in Plato's doctrine of ideas is a second-order idea. He also needs to consider the idea of the idea of ideas in his dialogues. But then he stops, not because he must, but simply because there is usually no explanatory advantage in going further. In the same way, we can find no explanatory soundness in going beyond the trope of precise similarity between two other tropes.[8]

Finally, it is worth noting that strict similarity is not a trope like others. To begin with, it is what we have called a *dependent* trope: it depends on the existence of things considered alike. Color-similarity, for instance, is an internal relation depending on the existence of color-tropes. Campbell

[8] As Anna-Sofia Maurin remarks, in a vicious infinite regress a considered statement (trigger) *is dependent* on the subsequent steps, while in a virtuous infinite regress, the subsequent steps depend on the considered statement, which makes them unnecessary (2007).

suggested that strict similarity is only a supervenient pseudo-addition that does not add any being to what already exists (1990: 37).

Nonetheless, if we take seriously our identifying condition for tropes, the fact that we are dealing with an internal relation does not make strict similarity or even higher-order strict similarities quasi-tropes or a non-tropes, as some theorists think. As already noted, the identity condition for the reality of similarities as tropes is satisfied, even if distinguishing strict similarity from other more primary kinds of tropes. If an essential condition for the existence of a (simple or not, homogeneous or not, external or not) trope is its spatiotemporal localizability, established by the application of its ascriptive predicative expression, we can argue that similarity is also spatiotemporal, though in a broad way. For example: when I consider the strict similarity between the colors of two shoes I see in a store window, this likeness would be somewhere in this place, which may include myself, but not in a distant place. My house and the Taj Mahal have a color-likeness: both are white. Nevertheless, I can swear that this likeness is situated on the planet Earth and not on the surface of the sun. Moreover, if my house or the Taj Mahal disappear, the color likeness also disappears, which means that the similarity also exists in time. Furthermore, when someone considers the similarity between the form of our Milky Way galaxy and the form of the Andromeda galaxy, this coarse-grained qualitative identity must have to do with the total distance between them, which is still localizable. But as great as this distance may be, it remains insignificant if compared with the immensity of the cosmos.

Problems for the theory of tropes do not stop here. What about other spatial relations? For example, the Golden Gate Bridge is (on the average) 67 m. above sea level. Certainly, this spatial relation is there and can even be measured. And this relation is located in space and time, enduring as long as the bridge exists and the average sea level does not change. This spatial relation isn't internal, insofar as it is independent of the relata only. This makes easier to classify it as a trope, but it is not because of this that it satisfies our identifying condition for tropes as spatiotemporally localizable entities.

But what about space and time in themselves? Normally we admit that only tropes and space-time exist. Even in realist ontologies, a separate existence of space and time was never seriously questioned. However, could space-time in some way consist of tropes or something derived from tropes? Imagine that all the world's objects and properties disappeared. Would space and time remain? We have the intuitive tendency to answer in the negative. However, according to a Newtonian theory of absolute time and space, the answer should be in the affirmative: space and time

would be individual-like entities. Space would be like a great container with material objects within it and would not cease to exist even if all the matter and energy ceased to exist and disappeared. On the other hand, according to the relational view originated from Leibniz, space could be constructed by means of relations, and this conception can easily be extended to include time. In the latter case, space and time could not exist in themselves, because being constructed of relations they require the existence of the *relata* (not necessarily material things). Both answers have always been controversial, and the discussion has been intensified by contemporary physics.

The attempt to explain absolute space and time in terms of tropes seems to be condemned to failure. If space as a whole is a trope, it cannot be located in space, and the same holds for time, contradicting our definition of tropes. However, it seems there is a good chance of explaining space and time relationally in terms of tropes if we begin with a modest commonsense approach. It seems clear that in primeval times people understood space by thinking of relations such as above, below, in front of, behind, inside and outside. We can localize an object *x* as being twice as far above object *y* as is object *z*. Originally time would also be relationally understood, by means of relations like earlier, present (simultaneous with the act of observation) and later. One can say that event *x* occurred three times as long ago as event *y* in relation to event z. Moreover, in order to make measurements, the plain man appealed to regularities as patterns: a *foot* to measure distances in feet, a *day* to measure periods of days… And one could with the aid of these regularities calculate speeds in order to conclude, for example, that Pheidippides could run more than 160,000 Greek Steps in one day before dying of exhaustion. This is how our usual concepts of space and time worked and still works in everyday life, where they do not demand a further explanation. The main point here is that all these relations should be tropes since they are also spatiotemporally located. However, since quality-tropes and material objects are also spatiotemporally located entities, it seems that we would end up in circularity: space and time would be defined as relations of spatiotemporally located property-tropes and objects as clusters of property tropes.

The answer to the circularity objection in this modest commonsense approach is that space and time are constituted by a network of spatiotemporal relations among spatiotemporal entities that can be quantitatively compared. For instance, consider the following rough description of the Southern Cross against the horizon: star *c* is seen twice as far below the smaller star *b* than *b* below star *a*, while stars *d* and *e* are

seen on opposite sides of *b* and (approximately) at the same distance from *b* as *a* is from *b*. With a similar approach, any particular spatiotemporal relation, for instance between *a* and *b*, could be located in the spatiotemporal network and because of this could be defined as a trope. And the same could be said of the individual star b as a spatiotemporally located cluster of tropes.

Of course, it is an entirely open question how such a rough commonsensical view could be developed, extended and transformed in order to comprehend the sophisticated and often controversial theories of contemporary physics. However, nothing could be more distant from the truth than to commit the naïve mistake of believing that the above account is so primitive and superficial that it could effortlessly be dismissed based on the discoveries made by modern science.[9]

3. Tropes and Concrete Particulars

The second major problem is that of constructing concrete individuals by means of tropes. For D. C. Williams, a material object is a set or sum of different conjoined tropes (1953: 11 f.). The advantage of this view is that it enables us to abandon the old, obscure concept of substance understood as some hidden substratum of properties. For the trope theorist, the material object turns out to be a kind of artichoke consisting only of its leaves, which are tropes.

The key-concept here is that of *compresence* (also called concurrence, togetherness, etc.), which can be understood as *the sameness or near-sameness of the spatiotemporal location of tropes.* The concept of compresence can easily be analyzed as composed of two other concepts: *co-location* and *co-temporality*. The co-location of tropes is their joint location in space, leaving aside *when* each of them comes to be located. Thus, two persons who take turns sleeping in the same bed can be said to be co-located in this place. The co-temporality of tropes is their simultaneous existence during the same time interval. Thus, my friend Magda and I are co-temporal, though not co-located, since we are very distant in space. The compresence of tropes arises only when they are co-located *and* co-temporal.

A naïve but instructive objection to the view according to which concrete objects are clusters of tropes is that if it is true, then all predication turns out to be tautological: the utterance ‘This chair is yellow’

[9] There is no *prima facie* reason to believe that a relational view cannot in principle be made compatible with general relativity theory.

would be tautological, because yellow is predicated of a subject that already has the trope yellow as a constituent (Loux 1998: 103). This objection is easy to refute. We just need to distinguish *necessary* from *contingent* tropes. As has been pointed out, a material object can be identified by means of an indexical added to a sortal predicate, as in the statement 'This is a chair' (Tugendhat 1983, Ch. 9).[10] Now, the necessary tropes are those typically specified in the definition of the sortal. Thus, 'a chair' is defined as a non-vehicular seat with a backrest, designed to be occupied by only one person at a time. The seat is constituted by one sub-cluster of tropes, the backrest by another, and the conditions that this complex object is non-vehicular and designed to be used by only one person at a time are constituted by dispositional tropes, variations and alternations of tropical relations that complete the definition. There are also contingent tropes, like those constituting the sub-clusters of armrests or four legs, since there are chairs without armrests and chairs without legs; and there are still more variable tropes associated with a chair, like its color, the relation to a certain person sitting on it, its distance from a table… The concept of a *chair* is one of an artifact. But we can consider natural kinds in a similar way. Gold is defined as an element with the atomic number 79, a dense, yellow, precious metal. However, its having a determinate atomic number is a necessary trope, though gold does not have to be yellow or even considered a precious metal, since these are contingent tropes.

Peter Simons gave a helpful answer to the question of the nature of material objects by pointing out that they should not be seen as an unstructured cluster of compresent tropes. A material object is typically made up of *a nuclear kernel of necessarily interdependent tropes giving a foundation to an accidental halo of contingent tropes.* The halo-tropes can be replaced by tropes of other kinds, but the kernel-tropes cannot (they can be approximated to sortal predicates). A consequence of Simons' view is that the halo-tropes are specifically founded on the kernel-tropes, while the kernel-tropes only generally found the halo-tropes (1994: 376 f.). Moreover, Simons accepts the possibility of variations: a concrete object formed only by kernel-tropes, etc.

Here a much more precise definition seems to be simply impossible. Stones, for instance, are material objects that can be composed of very different materials, having few tropes to individualize the object-kind

[10] Tugendhat defines a *sortal* as a predicate that has criteria for the spatial delimitation of the object, allowing us to distinguish what does or does not belong to it.

stone, with the exception of hardness, solidity, weight, volume, and color, all of them compresent. However, based on this cluster of properties, often combined with spatiotemporal determinations, we are already able to re-identify the stone as the same one.

Unhelpfully, compresence and kernel-tropes are still not enough to define material particulars. Socrates' wisdom is a dispositional property consisting of a very complex property-trope, as it seems. These tropes appear to have compresence, since they all seem to be located where Socrates is. Moreover, they could be individuated by a sortal predicate delimiting the spatiotemporal location of Socrates ('There comes Socrates again with his inconvenient wisdom!'). Finally, they can have a kernel: the 'peculiar core of the inconvenient Socratic wisdom.' But it is not a material object, not even an individual, insofar as it is said to *belong* to the individual Socrates and others could in principle, at least, share strictly similar qualities of Socratic wisdom. A common rainbow is constituted by co-located and co-temporal tropes of colors and forms – the seven colors of the spectrum – jointed together in a structured kernel, but it is less than a material object. The holographic projection of a teacup also has a proper compresent set of colors and forms. They belong to its kernel as an individual. But despite having colors, spatial extension, and form, it is no material object.

One strategy to deal with this problem is to add to the core of compresent tropes some tropes necessary for the identification of our typical material objects like:

volume,
form,
hardness or solidity (measured by resistance to pressure),
weight (depending on the presence of a gravitational field),
mobility in space…

This already excludes the property of Socratic wisdom and individuals like the rainbow and the holographic projection. But liquids, although they are material substances, do not have a specific form or solidity, unlike a stone, a tree or a table. For example, water takes the form of its container, and additional water can be added to a given quantity of water, increasing its volume. In a frozen state or as water vapor it ceases to be liquid. Resistance to pressure can be lower or higher. The water in a glass is already a material entity and an individual, though not properly a material object, since it lacks definite form, is not solid and has only limited resistance to pressure. A cloud has a low level of materiality: its droplets

have minimal resistance to pressure and it has no fixed and necessarily defined form. And what about supposed material entities like bacterias, viruses, atoms quarks or hypothetical super-strings?

My final condition is based on the already discussed assumption that our commitment to modest common sense does not exclude science.[11] We can refine the idea of hardness or resistance to pressure by proposing that a necessary trope constitutive of the core of any physical object is a derived trope that physicists call *inertial mass*. In physics, the inertial mass of a body is broadly defined as its inertial resistance to acceleration when forces are applied to it (an idea accepted in both Newton's and Einstein's mechanics[12]). This seems to me the most pertinent characteristic of what we call matter. Energy also has mass, but it isn't inertial mass.

I conclude that in an inevitably vague characterization, having the expected inertial mass, some size… and compresence of its definitional tropes would be necessary for singling out a material object. This excludes electromagnetic, gravitational, weak and strong forces, which are better seen as tropes. However, one cannot generalize this result to any individual. Consider the cases of a cloud, a rainbow, and a shadow. Consider the case of a crowd or the British Empire. These individuals do not form a material object or a physical body. Unlike material objects, a crowd and the British Empire are composed of tropes that are at least partially grounded on material, not tightly connected physical entities. And a historical tropical event like the Battle of Hastings was a spatio-temporal tropical event, not a material object. They are all complex structures made up of tropes, including mental tropes like intentional states and depending on material entities to be spatiotemporally located, even if only in a vague way. Since these tropical entities are independent and unequal and identified by nominal terms, they are individuals (Ch. IV, sec. 7).

A more technical difficulty arises from the alleged fact that the idea that particulars are clusters of tropes is vulnerable to a regression argument analogous to the third man argument used against the abstract objects assumed by a Platonist ontological view. Thus, suppose that a concrete particular were constituted only by the tropes T_1, T_2, and T_3. Since the

---

[11] J. L. Austin objected that terms like 'material object', 'material thing' and 'sense-data' do not originally belong to our ordinary language (Austin: 1962). Against this, we can only repeat that there are gaps left unexpressed by ordinary language, later filled by new philosophical terms (See Ch. II, sec. 6 of this book; see also Grice 1989: 227).

[12] As is well-known, the reason why according to relativity theory a body cannot reach the speed of light is that at this speed its mass would become infinite, requiring infinite force to accelerate it.

relation of concurrence could not be an abstract entity, it must be a trope. Call this relation Tc. In this case, it seems that we need a new concurrence for T_1, T_2, T_3, and Tc, which will be Tc', and so on infinitely (Daily 1997: 158).

My proposal to answer this objection takes a form similar to what realist philosophers have applied in defense of their own abstract properties. Compresence is made up of co-location plus co-temporality, which are spatiotemporal delimitations that remind us of the already considered cases of form and duration. They are all *dependent relational tropes* that must be considered *sui generis*, behaving somewhat like Platonic ideas with their resistance to self-predication. In other words: although you can meaningfully say that this red is red, and even that this triangle is triangular, you cannot meaningfully say that a concurrence is concurrent. Concurrence is a *sui generis* non-self-predicating limiting trope. Strict similarity is also a *sui generis* non-self-predicating dependent relational trope because one cannot say of the strict similarity between T_1 and T_2 that it is strictly similar without rising questions like: 'what would strict similarity be similar to?'

4. Formal Tropes

What should we say about formal entities like natural numbers? Numbers are often seen as Platonic or semi-Platonic universals. And they would not be tropes since they do not seem to be spatiotemporal. However, this isn't so uncontroversial! Much of our empirical world is made up of countable things. Would the number 3 exist if the world did not exist? Though this is an odd question, the tendency is to answer in the negative. For an empiricist like Locke, the number would be a primary quality (a trope), together with solidity, extension, figure, motion or rest, which are accessed by diverse senses and should remain the same independently of the perceiver (1690, Book II, Ch. VIII). Indeed, I can perceive one, two, even six things at a glance and these seem to be spatiotemporally located tropes; and some *savants* are able to perceive hundreds of things at a glance. To use an example borrowed from Penelope Maddy, it seems that the ten fingers of my two hands are in some way here (1990: 87). It seems that even a thousand grains of wheat scattered in the wind remain spatially and temporally located, though in a diffuse way. And if the insufferable rock band called 'The Fevers' flies from São Paulo to Rio de Janeiro, it seems that the number of their members has also moved. However, it is important to note that these trope-numbers are dependent on countable entities of our choice.

One can associate this dependency with Frege's account of numbers as properties of concepts, since as he has taught us, things to be counted must be first conceptualized. The question 'How many?' only makes sense if followed by a conceptual expression. For instance, if the concept is of the fingers of my hands, they are ten, but if the concept is of my hands, they are only two. The property of the concept of those grains of wheat scattered in the wind is that there are a thousand. And the movable property of being five is a property of the concept of *The Fevers*. Moreover, as Frege famously wrote, the attribution of existence is the negation of the number zero (1892, sec. 54).

So it seems that the concepts of number and existence are related. In fact, one can suggest that the property of existing and the property of being a number are higher-order tropical-properties because, like tropes, they are in a vague way spatiotemporally located: this black spot on the carpet exists here and now and not somewhere outside in a remote time. And it seems plausible that when I say 'This is my one and only nose,' 'these are my two hands,' 'these are my ten fingers,' the number one I am applying is located where my nose is, the number two is where my hands are, and the same with my ten fingers. The naïve error would be only to confuse these ethereal tropes with those qualities primarily constitutive of the nose, the hands, and the fingers. Indeed, numbers, as much as the existence of things, do not seem to be in outer space or in ancient times or in the solely intelligible realm of abstract ideas.

These considerations seem to be valid for applied arithmetic, insofar as numbers are first used to count empirical objects. After all, we learn numbers by counting material things: 'There are two apples and one pear in the basket, totaling three pieces of fruit.' In this case, the ascription rule of the predicate '…fruit in the basket' was applied to three distinct objects, attributing physical existence to each of them and showing in the process of counting that the rule has the higher-order trope-property of being applicable three times in an additive way.

In the view defended in this book (See Ch. IV) a concept is a rule, which means that the attribution of *existence* is here the second-order property (or trope) of a dispositional first-order conceptual rule (always understood as a trope) of being satisfied by at least one thing. And in a similar way, an applied natural number would be the second-order property (or trope) of a dispositional first-order conceptual rule (or trope) of being satisfied by means of an idealized counting procedure, where counting originally results from the distinguishable applications of a first

order conceptual rule to things like material objects or events or qualities attributing existence to them *n* times...[13]

Using Fregean devices it is easy to formalize this suggestion using only countable tropical applications of (tropical) concepts and the (tropical) concept of existence. The affirmation of the number 0 is the negation of existence.[14] Thus, using V in place of the conceptual expression 'moons of Venus,' we can symbolize the idea that there are 0 moons of Venus as ~∃x (Vx), saying that the conceptual rule expressed by V isn't applicable at all. Using E to symbolize the conceptual expression 'moons of Earth,' we can symbolize the idea that there is 1 moon of Earth as ∃x [Ex & (y) (Ey → y = x)]. Here E is applied only once. And using M to symbolize 'moons of Mars,' we can symbolize the idea that there are 2 moons of Mars as ∃x [(Mx) (My) & (x ≠ y) & (z) (Mz → (z = x) v (z = y))]. Here M is applied twice. It is the application of a tropical ascription rule for two and only two moons of Mars.

Above we considered first order conceptual tropes together with higher order existence tropes and applied numbers as higher order numerical counting-tropes. However, I think we can separate or abstract the numerical trope from these other concepts. We can do this by representing these tropes of countability by means of *localizable sets*. Thus, I propose that we can represent the 0 in 'the moons of Venus' as the located non-countability (non-applicability) of a concept symbolized by ~a. Instead of the 1 of 'the earth's moons' we can speak of a set that has as its only member a located higher-order applicability trope or {a}. Instead of the 2 of 'the Mars' moons,' we can speak of a set that has two located higher-order numerical tropes as members, as follows: {a, {a}}. In this way we can represent an applied number 3 by the localizable set {a, {a}, {{a}}} and so on. Note that this 3 has the right complexity by containing {a, {a}} (=2) and {a} (=1). But the fundamental point here is that we are explaining applied numbers by means of spatiotemporally localizable sets of countability-tropes and by convention the null set. The set of Mars moons numerical tropes is spatiotemporally located in our solar system and not in the Andromeda galaxy or in the origin of time. And such sets are not Platonic or sub-Platonic entities!

At this point, one can object that we have until now explained only natural numbers applicable to things. One could, however, instead point

[13] Of course, there are large numbers that are uncountable for us. But they remain at least *ideally* countable. And they can be seen as later extensions that can be calculated by means of symbolic manipulation alone.

[14] I argue for a higher-order view of existence in chapter IV, sec. 11 to 19.

out that what really matters is the number of abstract arithmetic, the *universal* independent of its satisfaction by countable material objects or events. The suggested construction has indeed this limitation since it represents only one number among many identical numbers. The natural number 3, formulated as {a, {a}, {{a}}}, is a triad and *not what is common to all triads*, namely, the abstract universal three, the three-in-itself. Indeed, the only way to represent what is common to all triads seems to be the appeal to a Russellian set of all sets of the same kind, which has its own shortcomings like the axiom of infinitude, overpopulating our world with an infinite number of objects.

However, I think that in the same way as we have constructed universal quality-tropes without appealing to abstract sets, we can also construct universal number-tropes without appealing to abstract sets. I think we can derive the *universal concept* of number, the number-in-itself, from our spatiotemporally located tropes of counting. As we have seen above, an applied number can be understood as a trope, since it is spatiotemporally localizable as a second-order property of a potential conceptual rule resulting from its at least ideally countable applications. Consequently, in order to account for the universal as a set of equinumerous sets of applied numbers, we can appeal again to our disjunctive model.

In this case, for instance, it is conceivable that the number 2 in itself would be a disjunction between a located dispositional higher order trope-set of countable applications used as a model (e.g., the number 2 in the statement 'I have 2 hands') *or* any other strictly similar (equinumerous) located set of countable tropical applicabilities. Now, in order to get the number 2 as the 'abstract universal,' the 'two-in-itself,' all we need is to apply to the separated set of tropical applicabilities the same procedure we have applied to get universals from our usual quality-tropes. For instance:

> Number 2 (*Df.*) = a located model set of tropes of countable applicabilities {a, {a}}*, *or*… any further located set of tropes of countable applicabilities strictly similar (equinumerous) to {a, {a}}*.

In this sense, the number as a universal (or 'abstract entity') can be defined as:

> The higher-order property of a conceptual rule of being a located set of tropes of (at least ideally) countable applicabilities taken as a model *or* of any higher-order located set of tropes of (at least ideally) countable applicabilities strictly similar (equinumerous) to the first one.

Note that such constructed universals remain empirical since they are higher-order disjunctive property-tropes that can be found scattered across our whole spatiotemporal world. This makes graspable why something abstract like mathematics applies to the empirical world.

Assuming a definition like that, we neither stumble over controversial infinite sets of objects (as in Russell's definition) or over pure sets (as in von Newmann's and Zermelo's definitions) nor remain unintentionally limited to particular instances or directly committed to any differentiating concrete feature (as in naïve empiricist views). The conclusion is that even the abstract world of arithmetic (hence, mathematics) is made up of some sort of thin higher-order tropes. Such tropes, like some others, would be situated at the peak of a building whose originating genetic-epistemic foundations are our more feasible perceptually given quality-tropes, so that numerical tropes that can be univocally named in this way can also be seen as dispersed over the world and able to be meta-predicatively designated. Finally, I would not be surprised if even logical properties were susceptible to similar treatment!

Now one could object: aren't such formal properties not too *thin* to be tropes? A dependent trope like a conceptual rule might be a thin trope. But a trope that is dependent on other possible dependent tropes will be still thinner so that formal tropes are simply too thin to be real tropes! However, isn't it a foolish prejudice to reject tropical properties only because of their thinness? There is no *quasi*-trope.

5. Conclusion

In this section, we have seen how trope theory can turn Platonic realism upside down. Much of what I have written here is speculative, still requiring a great deal of additional work and refinement. In this short space, I could do no more than offer a sketch of what seems the most consequent and plausible way to deal with the one-category ontology chosen to play a central role in this book.

CHAPTER IV

AN EXTRAVAGANT READING OF FREGEAN SEMANTICS

Wenn es eine Aufgabe der Philosophie ist, die Herrschaft des Wortes über den menschlichen Geist zu brechen, indem die Täuschungen aufdeckt, die durch den Sprachgebrauch über die Beziehungen der Begriffe oft fast unvermeidlich entstehen (…) so wird meine Begriffschrift, für diese Zwecke weiter ausgebildet, den Philosophen ein brauchbares Werkzeug werden können.
[If it is a task of philosophy to break the power of the word over the human spirit by laying bare the misconceptions that through the use of language often almost unavoidably arise … then my ideography, further developed for these purposes, can become a useful tool for the philosopher.]
—*Gottlob Frege*

…might the time not have come to reflect about the very foundations of analytic philosophy, and to see it as one task of philosophy to break the power of the mathematical sign over the philosophical mind?
—*Edward Kanterian*

The importance of Fregean semantics for the philosophy of language derives from its unique blend of theoretical simplicity, explanatory scope, and philosophical relevance. In this chapter, I want to revise and reconstruct the essentials of Fregean semantics. I intend to make it clear that his basic concept of sense can be paraphrased in terms of *semantic-cognitive rules* and that his concept of existence can be reconstructed in terms of the *effective applicability* of semantic-cognitive rules, leading to some unexpected consequences regarding the explanation of the concepts of verification, fact, and truth. With the identification of senses with rules, I intend to show the real link between Wittgenstein's semantics – that is, the way I understood his views in the last chapter – and Frege's semantics. This link was already noted by Michael Dummett, though he still offered no proper pragmatic exploration. Anyway, my aim here is not to produce a work of Fregean scholarship. It is instead to reconstruct Frege's semantic work with him, against him, and beyond him, in order to provide a more

rigorous framework for the rather vague semantic insights gained in the first chapters.

As is general knowledge, Frege explains reference (*Bedeutung*) using a semantic intermediary link which he called *sense* (*Sinn*) (1891:14). The schema below shows how Frege deals with these two main levels, (1) sense and (2) reference in the case of a predicative singular assertoric sentence (*Satz*) of the form *Fa*:

singular term: a	*general term: F*	*sentence: Fa*
1. sense	sense	thought
2. reference	concept (> object)	truth-value

Although Fregean semantics was a development of unparalleled importance for contemporary philosophy of language, it is not free from well-known oddities. My intuitively natural reading of its main semantic elements in terms of conceptual rules will also show how to purge Frege's semantics of its most puzzling eccentricities.

1. Reference of the singular term

Let's start with singular terms. The reference of a singular term is, for Frege, the object itself, taken in an enlarged sense. The reference of the name 'Moon', according to him, is the Moon itself with its craters. To designate the reference, he uses the German word 'Bedeutung,' whose literal translation in English is 'meaning.' Most English translators have chosen words like 'reference,' 'denotation,' and '*nominatum*,' in this way making clear what Frege really had in mind. There are also other terms, like 'semantic value,' 'semantic role' and 'truth-value potential.' These terms underline the contributions of the references of a sentence's components to the truth-value of the sentence as a whole. Although the literal translation of 'Bedeutung' as 'meaning' remains the correct one, for the sake of clarity I will use the word 'reference.'[1]

There is also an interpreter's discussion of the reason why Frege would have chosen the unexpected word 'Bedeutung' for the reference of a nominal term. A widespread interpretation is that one of the meanings of 'Bedeutung' (as well as of 'meaning' or 'signification') is *relevance* or *importance*, since reference is what matters most for truth (Tugendhat 1992: 231). While this may be the case, it seems clear to me that the strongest reason, at least with regard to the reference of natural language

[1] On the thorny issue of how to translate 'Bedeutung', see Beaney 1997: 36 f.

terms, is that by introducing the term 'Bedeutung' Frege substantivated the verb 'bedeuten.' In this way, the word no longer expresses the *act of pointing at* (*deuten*) or of *designating* (*bezeichnen*), but rather *what* is pointed at (*die Bedeutung*), *what* is designated (*das Bezeichnete*), that is, the reference itself.[2] These derivations could be diagrammed as follows:

Bedeutet... → *deutet...* *bezeichnet...* → *was gedeutet, bezeichnet wird/*
(means) (indicate... designates) (what is denoted, designated)

↓

die Bedeutung
(meaning = reference)

This would have been the small semantic twist with which Frege turned the word 'Bedeutung' into a technical term – a twist that seems to betray some semantic-referentialist influence.

2. Sense of the singular term

Now we come to what Frege understands as the sense of a singular term. To introduce it, compare the following two sentences:

1. The morning star has a dense atmosphere of CO_2.
2. The evening star has a dense atmosphere of CO_2.

Sentences (1) and (2) concern to the same thing regarding the planet Venus. But in spite of this, a person can know the truth of (1) without knowing the truth of (2) and vice versa. Frege's explanation for this is that although the two singular terms 'the morning star' and 'the evening star' refer to the same planet Venus, they convey *different informative contents*, that is, they have different *senses* (*Sinne*).[3] The word 'sense' is defined by Frege as *an object's way of being given* (*die Art des Gegebenseins des Gegenstandes*), which is usually translated as *a mode of presentation.* The senses of the singular terms 'the morning star' and 'the evening star' are

[2] Searching in the literature, the only place where I have found a similar view on this point is Kneale & Kneale 1985: 495.

[3] One can read singular terms like 'the morning star' as definite descriptions or as proper names (like 'The Morning Star'). I prefer to read them here as definite descriptions, since for proper names we can use the words 'Phosphorus.'

different, because 'the morning star' presents Venus as the brightest celestial body usually seen just before sunrise, while 'the evening star' presents the same planet Venus as the brightest celestial body usually seen shortly after sunset…

Frege writes that words *express their senses* (*drücken ihre Sinnen aus*), while senses *determine* (*bestimmen*) their reference, since the mode of presentation should show us how to find the reference. Even in cases where the reference does not exist, this determination of reference through sense is given as a possibility, since even in this case the words preserve their senses. This fact points to a flaw in Frege's idea that sense is the way an object presents itself to us, for in the case of empty terms there is no object to be presented to us. This is why sense can be better understood as the *intended* mode of presentation instead of as a mode of presentation given by the object (Textor 2010: 134); sense is the way we intentionally present an object or reference to ourselves, whether it exists or not. At any rate, for Frege an expression can have a sense without a reference, but cannot have a reference without its determination by means of a sense.

Frege extended his notion of sense to other terms and to sentences. In the case of the senses of (declarative) sentences, he calls it *cognitive* or (more literally) *epistemic value* (*Erkenntniswert*). The last term is also appropriate. The Fregean concept of sense has epistemological interest, for it constitutes the proper *informative content* of the linguistic expression. It is what makes 'the evening star' and other expressions informative. Or, using Dummett's words, 'sense is what we *understand* when we understand an expression' (1990: 92). The philosophical importance of Fregean semantics is largely due to the epistemological and ontological imports of the concept of sense (this is what distinguishes it from a more exclusively linguistic semantics like that of Ferdinand de Saussure.)

Frege is a Platonist about sense. For this reason, he conceives senses as abstract entities which can only be analyzed in terms of constituents that are also senses. A consequence of his Platonism of senses is that it prevents him from analyzing senses in terms of other concepts. However, it is just this task that naturally imposes itself. For it seems very plausible to understand senses as *semantic-cognitive criterial rules*. We see here a fundamental difference between Fregean semantics and the semantic considerations of the later Wittgenstein, who regarded senses or meanings as depending on episodic uses of expressions determined by rules. Dummett was perhaps the first to defend the idea that senses are rules as the most natural reading of Frege's use of the term senses. As he wrote in his book on Frege's philosophy of language:

> The sense of a word consists in a *rule* which, taken together with the rules constitutive of the senses of the other words, determines the condition for the truth of a sentence in which the word occurs. (1981b: 194; my italics)

And concerning the singular sentences in Frege, understanding with the term 'criterion' the condition of satisfaction of a semantic rule, he wrote:

> To know the sense of a proper name is to have a criterion for recognizing, for any given object, whether or not it is the bearer (referent) of that name; to know the sense of a predicate is to have a criterion for deciding, for any given object, whether or not the predicate applies to that object; and to know the sense of a relational expression is to have a criterion for deciding, given any two objects taken in a particular order, whether or not the relation it stands for holds between the first object and the second. (1981b: 229)[4]

The identification between senses and rules proves particularly compelling when we take numerical expressions as examples. Consider the following expressions:

$$1 + 1,$$
$$6/3,$$
$$(7 + 3) - 8,$$
$$(874 - 870)/2$$
$$5 - 3$$

All these numerical expressions have the same reference: the number 2. But their senses or modes of presentation are in each case different. At the same time, they are expressions of procedures, methods, semantic-cognitive rules or, in most cases, combinations of such rules by means of which we reach the identification of the same number 2 as a result (*See* Runggaldier 1985: 91 f.).

By treating senses as semantic-cognitive rules and these rules in the primary case as *shared conventions*, we contrast them with what Frege called *colorations* and *illuminations* (*Färbungen* and *Beleuchtungen*), which are feelings often associated with *image representations* (*Vorstellungen*) and *sensory-perceptions* (*Anschauungen*), as such all belonging to an intrinsically subjective level (Frege 1892: 31). These 'colorations' and 'illuminations'

[4] As shown in the introduction, Ernst Tugendhat later defended a similar understanding of the meanings of singular statements in a more systematic and detailed way, though refraining from doing it as a reconstruction of Frege's semantics.

are names for what we would more often call *expressive meanings*, that is, sensory-emotional states that we normally and customarily associate with expressions. Thus, for example, the words 'love,' 'dog' and 'hell' in the sentence 'Love is a dog from hell' (Bukowski) contrastively associate words linked with strong specific emotions in order to create a weird epigrammatic effect.

As Frege realized, the kind of appeal or lack of appeal that the colorations associated with words have for different persons depends correspondingly on similarities and differences in their human natures. Because of this, they do not require conventions to be communicated, as in the case of senses. This is why some people are emotionally moved by a certain poem, while others are not. Consequently, it is very difficult to translate poetry, which depends so much on colorations acquired by expressions in a particular language and form of life. Hence, colorations are not results of conventional rules; they are rather *regularities* originating from shared aspects of human nature within a historically developed cultural context. If my understanding of Wittgenstein's argument against private language is correct, then his attempt to explain phenomenological language as a simple replacement of public behavioral criteria like uttering 'ouch!' under conditions that would cause pain with a sentence like 'I feel pain' is insufficient (1984d, sec. 244). It is an attempt to assimilate the referential meaning of the phenomenal language to its expressive meaning (I suppose that both can be legitimated).

If in opposition to Frege we accept the view that sense is usually only something with the form of a rule (*etwas Regelartiges*), namely, a convention or a combination of conventions, we can easily solve the problem of the communicability of senses that has long tormented philosophers like him. This is because the reason can easily be found for the objectivity (interpersonal accessibility) of senses, as well as for their consequent communicability. This reason is that Fregean senses are epistemic unities easily reducible to conventional semantic-cognitive rules or associations of them, and such conventions are interpersonally established and agreed upon in a pre-reflexive manner. Indeed, accepting the conclusions reached through our discussion of Wittgenstein's views, senses typically result either from the direct application of interpersonally established semantic conventions or, more importantly, from associations or combinations of these conventions.

Accepting that the sense of a singular term is the same thing as a rule understood as a conventional or conventionally grounded procedure that plays a decisive role in the identification of the object, it is easy to go further and accept that this rule can be typically expressed by means of

definite descriptions. Hence, the sense or mode of presentation expressed by the singular term 'the morning star' is a conventional rule that can be understood as requiring as a criterial condition for the cognitive identification of the morning star that we see as the brightest celestial body not too far from the Sun just before or after the Sun rises. Concisely stated, this rule can be expressed by the definite description 'the brightest celestial body that is seen close to where the Sun is about to rise.' Without assuming that definite descriptions are expressions of rules, Frege also approached this in a note on the name 'Aristotle' (Frege 1892: 28). For him the proper name 'Aristotle' abbreviates a cluster of modes of presentation of the object that can be expressed by descriptions, which may include (i) 'the disciple of Plato,' (ii) 'the teacher of Alexander the Great,' and (iii) 'a person born in Stagira.' If this is the case, then (i), (ii) and (iii) express different senses, different rules that in one way or another help us to determine the reference of the proper name 'Aristotle' (*Cf.* also Frege 1918-19: 63).[5]

Of course, there is a controversy about this issue, which arose from Kripke's arguments against descriptivist views of proper names like Frege's. However, it seems to me out of question that Kripke's arguments are successfully countered by the kind of meta-descriptivist bundle theory suggested in the Appendix to Chapter I of the present book.[6]

3. Reference of a predicative expression

Frege has something to say about the reference of a predicative expression, which he calls a concept (*Begriff*) and which may include relations. This is odd because it seems natural to call a concept something like the *sense* of a

[5] If we compare these two passages, it becomes clear that in opposition to Kripke's interpretation (1980, Lecture I), Frege already had in mind the essentials of the later bundle theory of proper names. The same can be said of Russell (*Cf.* Russell 1911, Ch. 5).

[6] Assuming Kripke's views, François Recanati replaces senses with *mental files* as supposedly non-descriptive modes of presentation (2012: 34). However, since these files are clusters of information and not subjective *Vorstellungen*, they should be able to be linguistically expressed by means of descriptions, bringing us back to the descriptivist standpoint. For this reason, it seems that semantic-cognitive rules are able to do the same job with higher explanatory potential and (as we will see) with important epistemological consequences. Moreover, these rules or combinations of rules do not need to contain less information than files. They can be as informational, durable, transitory, changeable and flexible as required by the context.

conceptual expression – the mode of presentation of its *designata* – while the reference itself should be called a *property* (e.g., a red patch) or some *combination of properties* (e.g., a bird's colorful feathers).

A traditional philosopher like Kant understood the concept as immediately related to a *schema*, which, as I understand him, is a *rule* able to lead to the formation of a manifold variety of sensory patterns that are satisfied by those things to which the concept applies (*Cf.* Kant 1988, B 180). Although Kant's text on schematism is terminologically impenetrable, it is easy to paraphrase his intuition using the terminology we have developed based on Wittgenstein by saying that a concept is a semantic-cognitive rule or procedure that requires the satisfaction of criteria by particularized properties (p-properties) or tropes, which is also consonant with Dummett's and Tugendhat's analyses of singular statements. Consequently, we have good reasons to suspect that a concept should be the sense of a predicative expression, its mode of presentation, and not its reference, as in Frege's strange use of the term.

To be fair to Frege, he also says that when an object falls under a concept, the concept may be called a *property* (*Eigenschaft*) of the object (1892: 201), seemingly acknowledging that 'property' is the right term for the reference of a predicative expression. However, for him the criterion of identity for two concepts is the sameness of their value-range (*Wertverlauf*), what includes their extension, which means that predicative expressions with different senses but the same extension must refer to the same concept (2001: 31). So, for instance, '…animal with a kidney' and '…animal with a heart' should be predicative expressions referring to the same concept since they have the same extension. But it is intuitively obvious that kidneys and hearts are very different concepts.

In addition to belonging to the realm of reference, Frege also sees his concepts as *functions*. The mathematical concept of function can be defined as a rule that has as its input *arguments* and as its output *values* (for example: '$3 + x = y$' is a function by means of which when we give as input the number 2 as the argument for x, we get as an output the number 5 as the value of y). For Frege, a concept is a function whose argument is the object that 'falls under it' (*fällt unter etwas*) or does not and whose value is a truth-value, which can be alternatively two abstract objects: 'The True' (*das Wahre*) when the object falls under the given concept and 'The False' (*das Falsche*) when it does not. For example, the concept designated by the conceptual term '...is a satellite of the earth' has the value true for the object Moon and the value false for the object Jupiter.

Nevertheless, for Frege, concepts cannot be objects, either collections of objects, nor extensions (2001: 26). The reason is that objects,

collections of objects and extensions are *complete* (*vollständig*) entities. That is, they do not require anything to complete them. A concept, by contrast, as a function, is seen by Frege as necessarily open: he calls it an *incomplete* (*unvollständig*) or *unsaturated* (*ungesättigt*) entity, needing to be completed by those arguments represented by the objects falling under the concept. In contrast, objects referred to by proper names are complete (*vollständig*), saturated (*gesättigt*) or independent (*unabhängig*).

One could add that the saturated-unsaturated distinction can be found on three distinct levels: linguistic, semantic and referential. For instance: the predicate '…is a horse' could be called an unsaturated linguistic expression (the unsaturatedness is shown by the gap '…'), expressing a supposedly unsaturated sense, which refers to an unsaturated concept (property) as the ultimate unsaturated ground. This unsaturated concept, for its part, becomes saturated when some object falls under it, for instance, the object named 'Bucephalus' referred to by the predicative sentence 'Bucephalus is a horse.'

With metaphors like those of 'unsaturation' and 'incompleteness,' Frege hoped to open the way to the solution of the problem of the logical distinction between the subject and predicate of a sentence. After all, the subject (the singular term) would refer to the saturated object, which would complete the unsaturated concept referred to by the predicate (general term).

Unsaturated predicative expressions and saturated singular terms combine to form saturated singular sentences like 'Bucephalus is a horse,' which being complete must also be the name of an object, which for Frege is the truth-value of the sentence. This seems to be confirmed by the possibility we have of nominalizing sentences in the form of definite descriptions, since the latter are also singular terms (1879: § 3). Thus, the sentence 'Bucephalus is a horse' can be transformed in the description 'the horse named Bucephalus,' which appears in the sentence as 'The horse named Bucephalus was black.' The problem with this argument is that the same can also be done with general terms: '…is a horse' can be nominalized as 'the horse,' as found in sentences like 'The horse is an herbivorous animal.' Hence, this argument isn't persuasive. Anyway, we can accept that assertoric sentences are like proper names in the sense that they do not require completion as unities of meaning.

4. Ontological level

Discussing the unsaturated nature of the references of predicative expressions leads us to the question of the ontological nature of what

Frege meant by a concept. If a concept is an unsaturated entity, what kind of entity is it? If it is an abstract entity, it seems that we should also have concepts as referred-to abstract entities of empty predicates, like '…is a yeti,' which seems to be an ontologically abusive admission.

Anyway, it is by now clear that Frege uses the word 'concept' as a technical term that contrasts too strongly with the word's ordinary use. For our ordinary language intuition, there is surely an empty concept expressed by the predicate '…is a yeti,' but this concept should be called empty because it is nothing but the sense of a predicate that has no reference at all! It is no wonder that Frege has nothing to say about the sense of predicative expressions, since he has beforehand emptied them by absorbing the semantic level into the ontological one.

My final conclusion is that it is better to drop the Fregean technical notion of a 'concept.' This is a problematic remnant of ontological realism that does nothing to explain predication. Instead, I will understand the word 'concept' here in an intuitive way as the *sense* of the predicative expression: its mode of presentation of something. It is counter-intuitive to assume that '...is a yeti' must have any reference; but this predicate clearly has a sense intuitively expressing what we ordinarily understand by a concept, namely, that of the abominable snowman of the Himalayas. Thus, it seems that the best way to give a legitimate role to the word 'concept' is to see it as the sense of a predicative expression understood as its cognitive meaning, that is, its ascription rule.

5. Referring to particularized properties: trope theory

But if we drop Frege's technical notion of concept, what is the reference of a predicative expression? I think that nowadays the most reasonable answer to this question consists in an appeal to the pure ontology of tropes proposed in the Appendix of Chapter III of this book, since it not only promises a parsimonial solution for ontological problems, but produces less difficulties than the traditional doctrines. Thus, I propose to replace Frege's reference of predicative expressions with what we now call a trope, which I characterize simply as *any spatiotemporally individualizable property, notwithstanding its degree of vagueness*.

There are many examples of tropes that are genetically primary and directly accessible to experience: the white color I see when I look at newly fallen snow on a sunny day, and which is there in my visual field, the smooth surface of this couch, the rectangular shape of my computer screen, its hardness or my headache. All these are tropes – spatiotemporally particularized properties or simply p-properties – that

may range from simple objective or subjective qualities to complex ones, and from homogeneous or heterogeneous complex tropes, like the music I listen to in the former case and the personality of a human being or a country's political system or a social upheaval in that country in the latter, since all these things are in a less specific way also spatiotemporally localizable. Also very indirectly experienceable things like physical forces could be derivatively constructed from perceived tropes, since they are spatiotemporally localizable, and it is not inconceivable that even space and time, together with formal properties could eventually be reducible to tropes, as I tried to show in the Appendix of Chapter III.

Moreover, it is easy to suggest a particularistic construction of universals built on the basis of particularized properties or tropes. In my view, a universal can be disjunctively defined as:

Any chosen trope model T∗ or *any other trope strictly similar*[7] *to T∗*.

I suggest this assuming that the trope we take as the model T∗ is at our discretion and may vary according to the epistemic subject and even concerning the same epistemic subject on different occasions.[8] In this case, tropes T1, T2… Tn are identified as instantiations of the universal only because they are strictly similar (qualitatively identical) to an arbitrarily chosen trope model T∗. An additional point is that usually the trope-model needs to be intermediated by memory: we (usually) don't bring with us physical patterns to compare things with, but have a memory of them. The memory-trope cannot be the primary trope we intend to consider, since it must stand for the experienced one.

A material object could be constructed as a cluster of tropes. It can in principle be understood as a cluster of tropes displaying at least *compresence*, that is, it must consist of a co-located and co-temporal cluster of tightly connected varied tropes. Moreover, there are some general characterizing property-tropes like unity, displaceability, volume, solidity, resistance to pressure – scientifically explained in a broader way as inertial mass – that typically comprise material objects.

---

[7] Mere similarity would not do, as this concept is intransitive. Strict similarity means here the same as qualitative identity, which is transitive. Strict similarity must also be a trope, since it is spatiotemporally located between tropes, even if, as an internal relation, it is a subordinate trope.

[8] I suggested this disjunctive construction of universal by means of tropes as the best way to circumvent the usual but problematic definition of a universal as a set or sum of tropes that are strictly similar, one with the other (See Appendix of Chapter III).

I usually avoid using the word 'property' instead of 'trope', not because it isn't the best one, but because the philosophical tradition has too often hypostasized this word as referring to some scarcely intelligible non-empirical entity, vitiating our philosophical language. This tradition has stubbornly ignored the fact that in ordinary language the word 'property' has always been used to refer to simple or complex, homogeneous or heterogeneous tropes. Anyway, I intend to use the word trope exactly as the word 'property' is ordinarily used. Thus, I explicitly include among the tropes complex tropes made up of different kinds of tropes, these complex tropes possibly being designated by composite predicates like '…a black horse of the best Thessalonian strain' in the sentence 'Bucephalus was a black horse of the best Thessalonian Strain.' This does not make this complex trope (complex property) a singular material object, mainly because, as we will see later, a singular material object, taken as an individual, is seen as able to exist independently if compared with the trope to which it is tied (in a different possible world Alexander's beloved horse, Bucephalus, could still exist even if he were just a tired old nag).

According to the understanding of the reference of predicative terms that I am proposing, a predicative expression like '... is white' in the sentence 'The moon is white' does not refer to any Fregean concept. It primarily *ascribes, denotes*, *designates* (or *refers to*) a particularized property, namely, a trope, which is the whiteness of the Moon as normally seen by observers on the Earth. Secondarily but distinctively, however, the predicate '…is white' also *alludes to* (or *connotes*) the fact that this trope exemplifies the universal property of whiteness, here understood in the already explained particularist way as this same model trope that is being considered, or any other trope that is like it. Summarizing, a predicative expression has mainly a *twofold function*:

(A) *An ascriptive function*: that of *ascribing* or *denoting* the trope (property) belonging to the object referred to by the subject term,
(B) *An allusive function*: that of *alluding* to or *connoting* the denoted trope or any other tropes that would be strictly similar to the model-trope that could be considered by the speaker as designated by the predicative expression, building what might be called the *universal,* here understood in an ontologically unobjectionable particularist way.

The allusive function is subsidiary to the ascriptive function: to identify a trope you do not necessarily need to grasp its role as an instance of a

universal.[9] Better said, as particularized properties tropes have not only ontological, but also epistemic priority if compared with their role in the identification of universals.

Furthermore – opposing the overwhelming influence of the logicist tradition – we have, as a still more subsidiary element: (C) the *extension*. Although relevant, differently from (A) and (B), extension isn't primarily associated with predication. Extension doesn't even need to be implicitly considered in the act of predication! However, it can be derived from the application of the allusive function of the predicate plus additional information, allowing us to infer or even find: (C1) an extension of tropes as the set of tropes strictly similar to the trope in question and (C2) an extension of objects as a set of objects having tropes strictly similar to the trope in question. However, in both cases the extension is a further element that is usually an only vaguely inferred set.[10] As a rule, you do not need to take it into consideration to use a predicate ascriptively.

6. Difficulty with the concept of unsaturation

The main objection to the idea of incompleteness or unsaturation is that it fails to serve its main purpose, which is that of distinguishing a predicative expression from a nominative or singular term. Between the object referred to by the subject and the property designated by the predicate, there seems to be an important *functional asymmetry*: the nominative term always refers to its object and cannot properly take the place of a predicate; on the other hand, it seems that we can easily turn a predicate into a subject by means of *nominalization*.[11] For instance, 'Socrates' in the

[9] Here I agree with Keith Campbell, who has suggested an epistemic primacy of identification over the generalizing function (1990: 24-25).

[10] Even D. C. Williams portrayed things misleadingly here. For him 'Socrates is wise' (or any *Fa*) means 'The concurrence [togetherness] sum (Socrates) includes a trope that is a member of the *similarity set.*' (my italics, 1953: 11)

[11] There are several asymmetries. The most discussed is probably the asymmetry of subjects and predicates regarding negation: you can negate the predicate, but not the subject (nominal term) (Strawson 1971, Ch. 5). The answer seems to me clear. The negation of the predicate means the admission of the inapplicability of the ascription rule to the object identified by the identification rule. However, since the application of the ascription rule is dependent on the application of the identification rule, whenever you negate the application of the identification rule of the subject you also negate the applicability of the ascription rule and in this way the whole statement. Hence, it is impossible to negate the subject as the nominal term alone.

statement ‘Socrates is wise’ always refers to its object and cannot properly take the place of a predicate, while ‘… is wise’ can be nominalized as ‘wisdom’ in a statement like ‘Wisdom is a virtue.’ To make the point more convincing, consider the following sentences:

1. <A man who lived in Antiquity> was called Socrates.
2. <Wisdom> is a property of Socrates.
3. <Xantippe’s husband> is Socrates.
4. <There> is Socrates!

In these sentences, the name ‘Socrates’ at least seems to occupy a predicative position. However, this name clearly continues to be used logically as a proper name, since the true logical form of these sentences can be easily expressed, respectively by:

1. <*Socrates*> was a man who lived in Antiquity.
2. <*Socrates*> has the property of being wise.
3. <*Socrates*> is the husband of <*Xantippe*>.
4. <*Socrates*> is in < *that place*>![12]

One cannot effectively transform a singular term as such into a predicate, while predicates seem to be easily transformed by nominalization into singular terms. However, we can show that the nominalized predicate is, in fact, a disguised universal predication: the sentence ‘Wisdom is a virtue,’ for instance, could be analyzed as, ‘For any *x*, if *x* has wisdom then *x* is virtuous.’ However, the asymmetry returns at this deeper level, since we cannot analyze a proper nominal term (like ‘Socrates’) in the same way. The asymmetry suggests that subjects and predicates play different logical roles in sentences, which requires explanation. The question is: can the

[12] Notice that the demonstrative ‘that’ does not have here the function of a constituent of the identification rule of Socrates, but expresses the identification rule of a certain place. In indexical statements like ‘This is a daisy,’ the demonstrative ‘this’ expresses a one-foot identification rule, localizing a place in time, while the sortal ‘daisy’ is placed as part of the predicate ‘…is a daisy,’ expressing the ascription rule. It is different from ‘*This daisy* is yellow’, in which the sortal ‘daisy’ is the characterizing part of the identification rule, whose localizing part is given by the demonstrative ‘this.’ The logical form of the statement ‘This is Socrates’ is already revealed by the relational statement ‘<This spatiotemporal place> is where <Socrates> is located.’ (For the role of localization and characterization in identification rules, see Appendix to Chapter I, sec. 1.)

Fregean distinction between saturation and unsaturation really do anything to explain the difference?

At first glance, the answer is in the negative. Frege's distinction does not explain the difference between subject and predicate in a logical sense, because it is also possible to suggest that a singular term and, therefore, its sense and reference, is unsaturated or incomplete! After all, what is the difference between:

> [Bucephalus, Silver, Black Beauty, Fury… Pegasus] …is a horse.

And

> Bucephalus is... [black, strong, restless, swift… of the best Thessalonian strain]?

In the first case, the concept '…is a horse' is a function that according to Frege may have as an argument any object and as a value a resulting truth-value, which for the object *Bucephalus* is 'The True' and for the object *Alexander* is 'The False.' However, it makes just as much sense to apply the same reasoning to the second case. One can suggest that the nominal expression 'Bucephalus is…' refers to an object that is a function that may have as its argument any property designated by any predicative expression. If it is the property *white,* it has as a value 'The False,' and if it is the property *black,* it has 'The True' as its value, since we know that Bucephalus was a black horse. The undesirable conclusion is that in a singular predicative sentence both the general and the singular terms can be viewed as unsaturated in the sense of denoting functions that can be supplemented by a myriad of arguments able to bring in 'The True' or 'The False' as the resulting values!

7. Unsaturation as ontological dependence

Notwithstanding, I think that the metaphor of unsaturation is not exhausted so easily. In chemistry, a carbon compound is said to be unsaturated when it contains carbon-carbon bonds that can be broken by the addition of hydrogen atoms, which make it a saturated compound. The hydrogen atoms aren't said to be unsaturated. Isn't there a hint in the metaphor of an answer that was not sufficiently explored by Frege?

In what follows, I hope to offer a reading of the reference of a predicative expression in terms of tropes that enables us to make a useful paraphrase of the Fregean distinction between saturation and unsaturation.

This paraphrase is inspired by the Aristotelian *independence* definition of the individual as primary substance:

> All the other things are either said of the primary substances as subjects or in them as subjects. For example, animal is predicated of man and therefore also of the individual man; for were it predicated of none of the individual men it would not be predicated of man at all… Thus, all the other things are either said of the primary substances as subjects or in them as subjects. So, if the primary substances did not exit it would be impossible for any of the other things to exist. (1984, vol. 1, *Categories*, sec. 5)

That is, some things can exist apart, and some cannot, and the former are substances.

I am not here worried in questioning if there are substances, what they are and if they are ultimately able to exist apart. However, applied to individuals or material objects understood as (at least) clusters of tropes displaying compresence, the independence definition suggests that *the objects typified by material things exist in a manner* relatively independent *of their tropes in the composition of facts understood as tropical arrangements in the world.*[13] Moreover, I hold that the individual referred to as a subject is only independent *relatively* to its predicated trope-properties, because the relation of existential independence/dependence is here understood in a way restricted to the internal context of the fact represented by the statement.

In other words, my suggestion is that the true dichotomy distinguishing subject from predicate is *between independence* and *dependence*, terms only rarely used by Frege. Thus, what distinguishes the *designatum* of a predicative expression in the fundamental case of a predicative or relational statement is that this reference is a trope (simple or complex, homogeneous or heterogeneous) whose existence as part of the fact *depends on* a cluster of selected compresent tropes constituting the individual referred to by the singular term, which is independent relatively to that trope. It seems that this fragile distinction is what that really distinguishes the references of logical subjects. Here are some clear examples supporting this view:[14]

---

[13] Ignoring Frege's theses that the reference of a sentence is a truth-value and that a fact is a true thought, I will in the present context call the sentence's reference a *fact*. This choice will be justified in the sections 21 to 23 of this chapter.

[14] I take these examples from Mulligan *et al.* (1984: 300, 301 and 306), though their point wasn't the same.

> Mary's smile depends on Mary's existence.
> The car's skidding depends on the car's existence.
> The snubness of Socrates' nose depends on Socrates' existence.
> Amundsen's expedition to the South Pole depended on the existence of both Amundsen and the South Pole.

These examples also make it clear that we do not mean that the dependent tropes (like those of smile, skidding, snubness, expedition to South Pole…) could not exist independently of other individuals as clusters of compresent tropes, but that they could not exist as they are independently of the individual or individuals belonging to the fact represented by the respective statements. – Qualitatively identical tropes of smile, skidding, snubness… could obviously exist in the dependence of *other* individuals.
Concerning singular statements, my suggestion can be summarized as follows:

> In the constitution of a fact represented by a true singular (predicative or relational) statement, the trope ascribed by the predicative expression only exists in the *dependence* on the existence of the compresent trope-cluster constitutive of the object(s) referred to by the nominal term(s).

Hence, it is important to see that the considered existential tropical dependence is relative to the fact it is a constituent (*Cf.* section 23).

In trying to explore this view in more detail, we can begin by remembering Peter Simons' nuclear trope theory of material objects. According to this theory, individuals are in the standard case formed by *an essential nucleus or core of mutually founding tropes*, which is necessarily surrounded by a looser cluster of *accidental peripheral* tropes, so that these peripheral tropes require the nucleus of essential tropes for their existence (See Appendix to Chapter III, sec. 3). To this we should add, as already noted for the relevant case of material objects, that belonging to the nucleus are typically tropes like those of hardness, form, volume and resistance to pressure or solidity, a trope that in physics was better elaborated under the label of *inertial mass*, all of them related by the dependent trope of compresence.

Unfortunately, the issue is not so simple. As we saw in the Appendix of Chapter I, the *identification rule* of a proper name requires for its application sufficient and predominant satisfaction of at least one inclusive disjunction of the two fundamental description-rules belonging to it, which are the localizing and the characterizing rules (*Cf.* Appendix

to Chapter I). This identification rule, as we also saw, can be satisfied by an indeterminate range of independent criterial configurations, in other words, tropes or configurations of tropes. This means that what Simons understood as a necessary nucleus of mutually founding tropes *may change* regarding one only individual in different counter-factual situations. Already considered examples are the Aristotle born 500 years later in Rome in one possible world and the Aristotle who in another possible world was born with cerebral paralysis in Stagira in 283 BC, son of Nicomachus… and was unable because of his disorder to write any philosophy. Hence, the nucleus of mutually founding tropes may be different within limits established by the identification rule. Consequently, in the case of objects referred to by proper names there is *no* necessary condition *in re* – no unique real essence of the object – to be expected, but only a nominal essence given by its proper identification rule, even if grounded on verified regularities. Peripheral tropes, on their side, would be those referred to by our auxiliary descriptions like (i) 'the teacher of Alexander' and (ii) 'the founder of the Lyceum.' And it is clear that the tropes designated by relations like '…the teacher of…' and '…the founder of…' are dependent on the existence of individuals like 'Aristotle,' 'Alexander' and the 'Lyceum' in order to exist as components of the facts represented by statements (i) and (ii).

Searching for a simpler example, I will now consider the singular term 'this chair.' I regard this phrase as an indexical name. This indexical name has an *identification rule* made up of two interconnected fundamental description-rules: a contextually dependent localizing description-rule establishing a spatiotemporal location (by means of the demonstrative 'this' and some indicative gesture) *and* a characterizing description-rule (by means of the sortal 'chair'). This characterizing description-rule is simply the definition of a chair as a non-vehicular seat with a backrest made for only one person to sit on at a time. We can say that the complex criterion for the identification of chairs added to the spatiotemporal location is what in this case forms the indispensable nuclear structure of this *designatum*. Symptoms of this chair, such as its having four legs and two armrests, or its being made of wood, are peripheral combinations of tropes. Moreover, if I say 'This chair is green,' the trope of green (in the described fact) exists in dependence on the existence of a complex of compresent tropes that forms this chair and would not exist without their existence.

These considerations allow us to better understand the corresponding independence-dependence relation regarding the compresent core of tropes of an object satisfying its identification rule and its contingent peripheral

tropes. Consider, for example, the singular predicative sentence 'Bucephalus is swift.' The predicate '...is swift' in this sentence applies to a contingent trope that constitutes *swiftness*, whose existence is here fully dependent on the existence of an object, Bucephalus, which is constituted by some core of compresent tropes constitutive of a living material object. On the other hand, the same distinction also applies to properties linked to individuals that are not properly material objects. A rainbow, for instance, is an individual (a cluster of compresent tropes), though not properly a material object. But consider the dynamic fact described by the statement 'That rainbow is fading away.' The fading away of a rainbow is a process-trope whose existence is dependent on the existence of the rainbow in itself.

Consider now the true relational sentence 'Bucephalus belongs to Alexander.' Regarding this fact, the contingent relational complex trope of *belonging to* could not possibly be found if Bucephalus and Alexander didn't exist as independent individuals formed by nuclei of compresent tropes. That is, the proper existence of the relation '…belongs to…' is here indebted to the existence of two more stable essential nuclei of mutually founding tropes constituting the two objects Bucephalus and Alexander. These clusters of compresent tropes referred to by the names 'Bucephalus' and 'Alexander' are concrete psycho-physical individuals that certainly exist independently of the existence of the relatively contingent complex combinations of tropes constituting the trope of '…belongs to…' since to have ownership we need the previous existence of individuals having this particular relational property.

A problem arises when we have independent countable things or sortals designated by predicative expressions. So, consider once more our definition of a chair as a seat with a backrest made for only one person to sit on at a time. Suppose now that I point to the chair and say, (i) 'This chair has two armrests.' Since the tropical clusters constitutive of having two armrests do not belong to the definition that makes explicit the nucleus, its existence as something that the chair should be dependent on the chair's existence. However, the predicate '…has two armrests' exists in the independence of the object referred to by the subject 'this chair,' since they can be separated from the chair, differently from its color or size. The solution to this problem is simply to see the above logical analysis as incomplete. The right analysis must take roughly the form: (ii) '<This chair> has <its first armrest here> and <its second armrest there>, and they are two,' pointing to the armrests, where '*x* having *y* and *z*' is the main property-trope, which is dependent on this chair and its armrests.

A related problem arises when predicates denote sortals belonging to definitional cores. Suppose I say, (i) 'This chair has a backrest,' where

'…has a backrest' is the predicative expression. The problem is not only that having a backrest belongs intrinsically to the object referred to by the singular term, but that the backrest exists independently of the chair. One can saw the backrest and say things like 'This backrest is green,' using 'this backrest' to refer to an individual. To this case, I suggest a similar solution. A more complete analysis of the sentence (i) will be (ii) '<This backrest> belongs (intrinsically) to <this chair>', where 'x belongs intrinsically to y' means that it belongs definitionally to the sortal 'chair' used to characterize the located individual y.

Very complex tropes (homogeneous or heterogeneous, mixed or not) are also existentially dependent on the individuals to which they are bounded. Consider some examples:

(1) <Céline> had a strange personality.
(2) <India> has a democratic system.
(3) <The ancient Spartan State> was extremely militarized.
(4) <The Vienna Philharmonic Orchestra> played the 5th Symphony.
(5) <The Irish potato famine> was caused by <the late blight>.

None of these tropes could survive alone. They need to be attached to some localizable and characterizable individual to which they belong.

Finally, what about formal names and sentences? Consider the sentence 'Three is an odd number.' This sentence describes a mathematical fact. Considering here ideas about what confers existence, we can think the number three without thinking that it is also an odd number, or 'the number two or any multiple of two added to the number one', which is the definition of an odd number. But there is no 'being odd' independent of a number. Hence, the existence of oddness factually related to the existence of the number three is dependent on the number three that we are taking into consideration.

Consider now the statement 'Two is a natural number.' One could argue that to be a natural number belongs to the definition of two as a kind of *genus proximum*, although not essentially to the (here seen as incomplete) definition of two as its *differentia*. Maybe this *differentia* could be given by our already suggested understanding of applied natural numbers as higher-order tropical properties of actual or idealized counting belonging to an effectively applicable conceptual rule (See sec. 4 of the Appendix of Chapter III). Repeating what I said there, consider the statement 'This hat has three corners.' Here the applied number 3 indicates that the possible conceptual rule identifying the corners of this hat not only has the tropical meta-property of being applicable (attributing existence),

but also the tropical meta-property of being applicable three times in an additive way (a counting process). Moreover, we can analytically express this conceptually dependent higher-order trope of 3 by means of the set of applications {a, {a}, {{a}}} understood as a spatiotemporally located higher-order numerical set-trope.

But how to represent the number 3 distinguishing it as the universal object that is common to *all* conceptual identifications of three singular entities, the three-in-itself? Here, if we wish to avoid speaking of a Russellian abstract set of all sets of the same kind, we can still construct the number 3 as a located model of tropical numerable trope-set {a, {a}, {{a}}} *or* any other strictly (equinumerous) located trope-set:

> Number 3 in itself (*Df.*) = a chosen higher-order located numerical set-trope of counting {a, {a}, {{a}}}∗ used as a model or any other higher-order strictly similar located numerical set-trope.

This definition still allows the predicate '…is a natural number' to be ascribed to the whole *definiens* as an internal dependent addition (a *genus*) and the predicate '…is an odd number' as an external dependent addition. In any case, even the name of a so-called abstract object, such as 'the number three in itself' cannot be moved to the predicate position here, insofar as it refers to something held as independent, being identifiable (existing) independently of its non-definitional predicates like '…is an odd number.'

Understanding unsaturatedness as *relative existential dependence* suggests, therefore, that the tropes denoted by the predicate have an inevitable tie of dependence when considered in relation to the relevant individual within the fact referred to by the singular sentence. This gives us a better understanding of the asymmetrical tie between subject and predicate.

Summarizing the argument, my point is that the independence/dependence distinction gives a sufficiently reasonable ontological ground (I guess the only one) to explain the logical distinction between the references of subject and predicate in singular predicative and relational sentences. The nominal term cannot be moved to the predicate position because it refers to a core of compresent tropes that exists in relative independence of the less central tropes in and outside of the core, these less central tropes being able to be designated by predicative expressions.

In my view, the proposed analysis also sheds light on the so-called problem of the unity of proposition. What really differentiates subject from predicate regarding the fact represented by the statement is the corresponding independence/dependence of their references. Moreover,

what assures the unity of the thought-content expressed by the sentence is simply the existential dependence/independence in the factual unity (for instance, in the fact that Bucephalus is swift). And it is clear that these ties of dependence/independence will be more evident when the difference in relevance between the elements in question regarding the identity of the individuals is greater, and weaker when this difference is smaller, justifying occasional uncertainties.

Finally, one could object that what really distinguishes the predicate from the subject in singular statements is simply that the subject is a singular term that identifies *one particular object and distinguishes it from all others*, while the predicate is a general term *able to be applied to more than one object*… It is this possible one-to-many relation that is at the base of the subject-predicate distinction!

Nonetheless, although this is true regarding a formal definition of singular and general terms, I believe that what gives a reason for this distinction is the relation of independence/dependence between subject and predicate. What defines an individual is that because of the uniqueness of its existence it can be referred to by a nominal term by satisfying its condition of sufficiency. And what defines a property-trope is its existential dependence on some individual (object). The individual is by definition non-repeatable. On the other hand, the property-trope is repeatable, insofar as qualitatively the same property-trope can be, by its lack of existential dependence, tied to many individuals. But this is so as a consequence of the fact that the existence of the property-trope must always be dependent on the existence of individuals, disregarding what individuals. In the end, it is the difference in nature between individuals (objects) and property-tropes (attributes) that is responsible for the one-to-many relation.

8. Sense of a predicative term

The independence/dependence relationship originating on the ontological level of reference is reflected on the semantic and linguistic levels. It is first reflected on the semantic-epistemic level of Fregean senses. We see this in the fact that the identification rule of the nominal term – its sense – is applied to its object independently of the ascription of tropes to the same object by the ascription rule – the sense – of the predicative expression, while the ascription rule of the predicative expression – its sense – depends on the prior application of the identification rule of the object referred to by the nominal term. Finally, on the level of linguistic signs,

the same relation of independence/dependence is what makes the singular predicative sentence take its usual subject-predicate form.

Our view of tropes as the *designata* of predicative expressions allows us to make some additions not present in Frege's original semantic distinctions. The first is the suggestion that different predicative expressions with the same *designata* may be able to have different senses, paralleling the case of nominal terms like definite descriptions. Consider the following two sentences:

1. *Mont Blanc* is white.
2. *Mont Blanc* reflects all wavelengths of the visible spectrum.

The reference of the predicative expressions of sentences (1) and (2) – the trope or compositions of tropes that constitute the whiteness of *Mont Blanc* – remains the same, while the senses of the predicative expressions are different: a person may know that *Mont Blanc* is white without knowing that its surface reflects all wavelengths of the visible spectrum and vice versa. This means that there are differences in concepts as modes of presentation or ascription rules of the predicative expressions of sentences (1) and (2), although they have the same *designatum*.

Another consequence of our understanding of predicative expressions as basically referring to tropes by means of their semantic-cognitive conceptual rules contradicts the Fregean expectation that the same sense cannot have more than one reference, which favors the potential for multi-referentiality inherent to predication. Consider the following sentences:

1. The South Pole is white.
2. *Mont Blanc* is white.

The predicate '...is white' in sentences (1) and (2) obviously has the same sense in both, as in each case it expresses qualitatively identical ascription rules. But the tropes of whiteness (of reflecting the combined wavelengths of the visible spectrum) of the South Pole are located at the South Pole itself, while the tropes of whiteness of *Mont Blanc* are located in its eternal snows. The same can be found in the application of relational predicates. This is explained by the fact that the different objects referred to by different singular terms have numerically different tropical configurations that satisfy qualitatively identical ascription rules of the same predicative expression.

9. Dependence of the predicative sense

As we have already noted, in the context of a singular predicative sentence, the identification rule of the singular term applies to the object as some core of compresent tropes, which seen as a whole exists independently in relation to its more or less dependent partial or peripheral tropes. Consequently, the identification rule is also able to be applied regardless of the application of contingent ascription rules, which means that this identification rule can be conceived as being applied in isolation. This explains its independence and why one could call it complete or saturated. The ascription rule, on its side, will be applied to a trope dependent on the core and consequently depending for its real application on the earlier application of the identification rule, lacking in this sense completeness. This is at most clear in the case of rules for contingent properties, like the conceptual rule for the predicate 'swift' when applied to Bucephalus.

The same may also hold for the fundamental descriptions constitutive of the identification rule of the nominal term in the sentential context. Since the tropes belonging to the object to which the identification rule applies are ultimately dependent on the existence of this object as containing a kernel of tropes, even the ascription rules of predicative expressions already belonging to the identification rule of the object as part of this rule require prior application of the whole identification rule to identify the object in order to become themselves applicable as part of the identification (*e.g.* the statement 'Aristotle was the author of the Metaphysics'). Because of this, the application of the predicate's ascription rule is always dependent on the application of the identification rule of the singular term.[15]

The general sense of a concept-word, which (diverging from Frege) we identify with the concept or ascriptive rule expressed by it, should then be a rule whose application to an object depends on the prior application of another rule. Hence, the ascription rule of the predicative expression is dependent, incomplete, unsaturated, in the sense that it demands the prior application of the identification rule of the singular term in order to be applied. It is necessary to identify, that is, in the empirical case to find some particular object in space and time, in order to be able to characterize it by ascribing the predicative rule to its appropriate trope. We must, for

[15] As Ernst Tugendhat wrote: ''Fa' is just the case to the extent that the rule of identification for 'a' is followed, and based on this result 'F,' is applicable in accordance with its rule of application'. (Tugendhat & Wolf 1983: 235)

instance, first apply the rule that allows us to spatiotemporally locate the horse called Bucephalus in order to apply to it related tropes, and on that basis, the ascription rules of predicative terms. Thus, due to the independence of the object Bucephalus, we apply the ascription rules for the predicates '... is a horse,' '... is black,' '... is swift'… and also the ascription rules of more complex predicative expressions like '…a horse that belonged to the best Thessalonian breed' to the tropical kernel constitutive of Bucephalus. And we also need first to apply the identification rules for Bucephalus and Alexander in order to be able to apply the relational predicate '…belongs to…' In a similar way, we need to apply the rule that allows us to mentally identify the number 3, in order to be able to apply to associated dependent tropes the ascription rules of predicative expressions like '…is odd,' '…is a prime number,' though it is not the case that the number 3 depends on these things in order to be identified as such. In the same way, the relational ascription rule for '3 < 7' is only applicable in dependence on the independent application of the identification rules for the numbers 3 and 7.

As I have very early noted (Ch. I, sec. 1), it would be a naive objection to think that after all it is possible to say things like 'That is a horse' or 'There is a black thing,' applying ascription rules of predicates without identifying Bucephalus. The reason is that a fully detailed identification of the reference as Bucephalus isn't required at all. Indexicals such as 'that' and 'there' accompanied by some gesture of pointing are already able to identify some spatiotemporally localizable spot which exists independently of further predication, being therefore in a technical sense an object or individual. As we already saw, this relative independency of the indexical identification rule can be made explicit when the indexical is followed by a term designating countable things, that is, a *sortal*, such as 'that object,' 'that animal,' since we localize with the demonstrative and characterize with the sortal. Therefore, not only does the trope designated by the predicate depend upon the previous existence of the object and its identification, but, as a consequence, also the effective applicability of the ascription rule of the predicate must be dependent upon the prior application of the identification rule to the relatively independent cluster of tropes. This is how the relation of semantic dependency – on the level of sense – mirrors the relation of ontological dependency – on the level of reference – solving the riddle of unsaturation.

10. The concept horse paradox

We can continue to make major revisions of Frege's views in order to overcome difficulties arising from his semantic views, like the so-called concept horse paradox. Based on his view of a concept as the unsaturated reference of a predicate, Frege was led to the strange conclusion that one cannot name a concept. For him the sentence:

1. The concept horse is *not* a concept,

is *true*. After all, 'the concept horse' appears here as a singular term – a definite description – and as such it must refer to something saturated, that is, an object and not a concept. The paradoxical point is that the denial of the true sentence (1), which is:

2. The concept horse is a concept,

must for Frege be *false*! Nonetheless, (2) clearly sounds like an obviously true analytic sentence.

From our perspective, the first thing to do is to treat nominalization as what it really is: an abbreviated way to speak about quantified concepts. What (1) really means is:

3. For any *x*, if *x* is a concept horse, then *x* isn't a concept,

which is *obviously false.* Regarding sentence (2) it really means:

4. For any *x*, if *x* is a concept horse, then *x* is a concept,

which is *obviously true.* Using H to replace '… is a concept horse,' which is the ascription rule able to designate the property-trope of horseness, and replacing '…is a concept' with C, which is the ascription rule able to designate any property-trope in an undifferentiated way, we can formalize (3) as (5): $(x)\ (Hx \rightarrow \sim Cx)$, which is false, and (4) as (6): $(x)\ (Hx \rightarrow Cx)$, which is true.

What is the lesson of this analysis? If 'the concept horse' does not really work as a definite description – as a singular term – but rather as a hidden universal predication, Frege was wrong in maintaining that it cannot be a concept only because it now works as a definite description. Frege's 'paradox' results from an incomplete analysis of sentences like (1) and (2) and the true analyzed sentences are the corresponding harmless

universal conditionals (3) and (4), the first being contradictorily false and the second tautologically true. If we agree that rightly analyzed 'the concept horse' expresses a universal predication and no real singular term, the whole paradox dissolves. It turns out to originate from the naïve mistake of thinking that if you put a predicate in the position of a subject, transforming it into a definite description, you necessarily transform it into a real singular term (See Appendix to this chapter).

11. Existence as a property of concepts

At this point, we can turn to Frege's treatment of the concept of existence. Deepening an idea already present in Kant's philosophy, he suggested that existence is a property (*Eigenschaft*) of a concept, namely, the property that at least one object would fall under it (Frege 1884, sec. 53). A similar idea was later advocated by Bertrand Russell in the suggestion that existence is the property of a propositional function of being true for at least one instance (1994: 232-3, 250-54.).

Here I will not try to interpret the details of Frege's often obscure remarks. Using more current terminology, I will follow an explanation taken from John Searle, who with his usual clarity brings us unmistakably to the point (2008: 176). Consider the sentence 'Horses exist.' This sentence can be analyzed as:

There is at least one ... such that (... is a horse).

As Searle notes, this sentence contains two components. One is expressed by the predicate '…is a horse,' symbolically H*x* (where we use *x* instead of '…' and H replaces 'is a horse'). The other component is the predication of existence expressed by the open sentence 'there is at least one ... such that ...' This predication can be symbolically expressed as $\exists x(...)$ (where $\exists x$ replaces 'there is at least one … such that…,' and the last '...' is the gap to be filled by *some concept applied to something*, now in the most proper ususal sense of the word concept, which in this case is the concept horse symbolized as H*x*. The result is that the whole sentence 'Horses exist' can be symbolized as $\exists x(Hx)$. This also means that the predication of existence $\exists x(...)$ is a metapredication expressing a higher-order concept, a concept of a concept, a metaconcept under which other concepts can fall – in this case (H*x*). Thus, $\exists x(Hx)$ instantiates the general form $\exists x(\Phi x)$, which usually expresses a second-order concept – the concept of existence – applied to some first-order concept. In a Fregean way of speaking, what this second-order concept does is to say of the first-order concept that at least one

object falls under it, which also means that the first-order concept is satisfied or fulfilled by being applicable to at least one thing. So understood, existence is something objective, since this satisfaction is independent of our cognitively grasping it as the *applicability* (and not mere occasional application) of a concept.

12. Existence as a property of conceptual rules

These last ways of speaking are more interesting to me because they could be paraphrased in accordance with my identification of concepts with senses of predicates, more precisely, with conceptual, semantic-cognitive ascription rules. This identification shows that existence can be a property of these conceptual rules, namely, their property of being able to be satisfied, fulfilled, or simply applicable. For instance, when I say 'Horses exist,' I mean that the conceptual rule expressed by the predicate '…is a horse' is definitely applicable. More precisely, I mean that this conceptual or ascriptive rule is, if it is given, *effectively applicable in a domain of external objects.* I add the adverb 'effectively' or 'definitely' to make it clear that I do not use the word 'applicable' in a merely subjunctive sense, as referring to something that *may* be applied, but as referring to something that *is effectively* (*definitely*, *warrantedly*) *applicable*, which is continuously the case during some period of time (the period in which the object is said to exist). Moreover, the own ascription rule must be seen as a *possibility*, not as an actuality, since things exist in the independence of their semantic-conceptual characterizing's rule existence. Furthermore, the existence or effective applicability of a semantic-cognitive rule is always considered with regard to a certain *domain of entities* (a 'universe of discourse'). The most fundamental domain is that of the *real empirical world*, be it the external (physical) world (Carnap's thing-world) or the internal (psychological) world. The statement 'Horses exist' applies in the first domain. The statement 'Headaches exist' applies in the second domain. Indeed, what is normally meant by the predication of existence isn't the applicability of a possibly given ascription rule of a general term as a mere possibility entertained only in our imagination, but also an effective applicability of the rule within some empirically given domain of entities. Furthermore, this effective applicability is usually within what we might call its *most proper* domain of entities, which in the case of horses is a domain of external, physical objects, and in the case of headaches is a domain of internal, psychological states. I consider this point here because

there are subsidiary cases, like that of the Valkyries,[16] whose most proper domain is mythical – that of Norse mythology.

As one could guess from the last example, there are other higher-order domains and sub-domains of entities within which we can predicate existence, even if only in a subsidiary sense. One can say, for instance, that Valkyries' horses exist in the fictional domain of Wagner's opera *The Valkyrie* in the sense that the ascription rules for these fictional horses are effectively applicable in the fictional domain described in the *libretto*. There are also cases like the probable existence of life in other galaxies, which can be in principle verified. Thus, there are imaginary mythological domains, fictional domains in the arts, and domains of imaginable but also plausible entities. Moreover, there are domains of so-called abstract entities and their various sub-domains, like the domain of mathematical entities, of logical entities… It is simply a linguistic fact that we can apply the word 'existence' in any of these domains. What I intend to show in the following is that there is a unifying justification for this.

According to the view I am supporting, to say that horses, rocks, trees and chairs exist is to confer effective applicability to the ascription rules of the respective concept-words 'horse,' 'rock,' 'tree' and 'chair' in the fundamental domain of material objects belonging to the objectively real external world. To say that thoughts, joys and pains exist is to ascribe effective applicability to the ascription rules expressed by the concept-words 'thought,' 'joy' and 'pain' in the subjectively real mental domain of entities. And to say that 'totalitarianism,' 'corruption' and 'exploitation' exist is to affirm the effective applicability of the ascription rules of these concept-words within the psycho-physical domain of social entities. The domain of entities to which such concept-words apply is usually assumed to be respectively physical, psychological and social. As a general rule, to say that an entity exists is to say that its conceptual rule is effectively applicable in the already conventionally established *most proper* domain of application. Thus, to give examples, the most proper domain of application of the conceptual word 'horse' is the real external world, while the most proper domain of application of 'Valkyrie' is a fictional one. That is, it is normally assumed that the attribution of existence is made in its most proper domain. But this assumption isn't necessary (when I say that there are horses in Wagner's opera *The Valkyrie*, the concept horse isn't being applied in its most proper domain).

[16] The Valkyries were maidens who served the god Odin, choosing the soldiers on battlefields worthy of admission to Valhalla.

As already noted, a concept – understood as the semantic-cognitive ascription rule of a predicative expression – is able to generate dependent, subjective criterial configurations. Thus, to say that a concept-word is effectively applicable is to say that dependent criterial configurations generated by its ascription rule are able to be fulfilled by corresponding independent, objective criterial configurations. These objective criterial configurations (external or not) can be seen as configurations of tropes usually belonging to more complex tropical arrangements called facts – another point against Frege that I will explain and justify in some detail later.

The parallel between the concept of existence in Frege and the more detailed concept of existence derived from my reconstruction of concepts as senses of predicates understood as ascription rules is straightforward:

> *Concept of existence (Frege) =*
> A second-order concept that demands for its satisfaction that a first-order concept has at least one object that falls under it.
>
> *Concept of existence (reconstructed) =*
> A conceivable higher-order semantic-cognitive conceptual rule that has a criterion for its (effective) application that a possibly given lower-order conceptual (or ascriptive) semantic-cognitive rule is effectively applicable to at least one entity, this entity being a trope or a configuration of tropes, usually in what is conventionally viewed as its most proper domain.

In my judgment, the advantage of this last form of analysis is epistemological: we are better able to scrutinize the nature of our existence-assignments, as will be shown by the answers to objections.

13. Two naive objections

There are two naïve objections to the proposed formulation of the higher-order view of existence, which offer revealing answers. The first is that the concept of a rule's effective applicability would be an anthropomorphic one, while things are said to exist in full independence of cognitive Beings.

However, this objection only arises if we confuse the concept of *effective applicability* (within a certain domain) with the concept of effective *application*. The application of a semantic-cognitive rule is an act or a series of acts that are essentially mental, though often also inevitably

sensorimotor, resulting in judgments. The application of the conceptual rule for the identification of the planet Venus, for instance, really demands the existence of cognitive Beings able to perform the application. Our judgment that the Moon circles the Earth depends on the experience of the application of a verifiability rule for the existence of this fact by ourselves or by someone who testifies to its application. On the other hand, the concept of effective applicability of a possible rule is not anthropomorphic. Even if there were no cognitive Beings able to apply the identification rule for the concept Venus, this planet would continue to exist, since if the ascription rule for the identification of Venus existed, it would still be effectively applicable to this object in its proper domain. The rule would still be applicable, even if no one had ever applied or even conceived it! The rule would be effectively applicable in a universe without any cognitive being able to conceive it, since all that is required is that *if* the rule existed, it would be effectively applicable. Thus, there is no doubt that the concept of effective applicability, as I understand it, isn't anthropomorphic.

This answer makes it easier to refute a second naïve objection. This objection could easily be made by proponents of the idea that existence is a property of things instead of concepts. According to it, if existence is a property of conceptual rules, then it has nothing to do with the objects that fall under these concepts: existence seems to be something floating above things that are said to exist. However, this seems odd, since intuitively we think that existence must in some way *belong* to entities that we believe exist!

The answer to this objection is that there is no contradiction between being a higher-order property of an entity and belonging to this entity. We make this clear by inverting the form of exposition. We can not only say that some possible ascription rules have the property of being effectively applicable to tropical properties belonging to a certain domain, but we can also say *that some tropical properties of a domain, the real ones, have the property of having their own ascription rules effectively applicable to them*, meaning by this that these entities *exist* in their most proper domain. That is, when we say that kinds of objects such as horses exist, we also mean that at least one of these conceivable countable kinds of objects has the *higher-order property or trope of having its ascription rule effectively applicable to it.* In other words, we mean that at least one horse has the meta-property of existing in the actual external world as part of it, and that this meta-property *is also a property of the kind of animal* – even if of a second-order – since it is a property-property at the level of the object's ascription rule, belonging to the object but not intrinsic to it.

In still other words, according to the higher-order view of existence, the red trope of a couch in front of me exists only insofar as this object (the couch) has the property of *falling under* the concept of being red in the Fregean way of speaking. But in a more natural way, we can say that the trope of redness of the indicated couch exists in the sense that the ascription rule of the concept-word 'red' has the meta-property of being effectively applicable to the couch's trope of redness. Even in a world where this ascription rule does not instantiate, for instance, a world without cognitive Beings to think and apply the rule, this rule (understood as *possibilia*) would remain effectively applicable, because we know that if this rule were conceived, it would be effectively applicable. (Since the rule only instantiates in minds, the rule is in this case only a possibility; but even if the rule actually does not exist, *the effective applicability of the possible rule* actually exists as a higher-order dispositional trope). However, this also means that the couch's trope of redness secondarily owns the meta-property of the effective applicability of its ascription rule to it – it owns this property-property dispositionally. That is, since the property of existence is the ascription rule's property of being effectively applicable to the trope of redness located on the surface of the couch, that property of the ascription rule is a meta-property of this trope of redness. It is so because, through the ascription rule, this property belongs indirectly but dispositionally to the trope of redness belonging to the real empirical world. Finally, the higher-order property of existence or effective applicability of the rule must be where the rule is, that is, it must be spatiotemporally located, being, therefore, a trope. Existence doesn't need to be an exception to our all-embracing trope ontology.

Considering that the meta-rule of existence is a trope that also applies to the trope, even if in a subordinate way, one could still ask: how would it be possible in the case of a possible world where there is no cognitive being able to think this cognitive-semantic rule? The answer is: the property of having the rule effectively applicable to it is a *dispositional* and not an actual property. In a similar way as an object is only dispositionally green at night when colors cannot be seen, the existence of an object will remain as a disposition, independently of the existence of cognitive beings able to identify existence by the application of conceptual rules.

Summarizing: it is a peculiar feature of the concept of existence (and certainly of some other concepts) that, being owned by a first-order concept effectively applicable to some entity, it must also be owned by some entity belonging to the chosen domain of entities without being a proper constituent of this entity.

14. Existence attributed to objects

The idea that existence is a property of concepts concerns not only what is meant by general terms, but also by singular terms, since both kinds of terms express conceptual senses, and their references can be said to exist. Since singular terms can be generally divided into proper names, definite descriptions, and indexicals, I will briefly consider each of them, beginning with definite descriptions.

Consider as an example the following definite description: 'the inventor of the Maieutic'. Applying the logical device to treat some descriptions by replacing them with a predicate, we symbolize the predicate 'the inventor of Maieutic' with M, so that the statement 'The inventor of the Maieutic existed' can be analyzed as:

$$\exists x\,[\mathrm{Mx} \,\&\, (\mathrm{y})\,(\mathrm{My} \rightarrow \mathrm{y} = \mathrm{x})].^{17}$$

In this way, we are affirming the existence of at least one and not more than one inventor of the Maieutic. This means that the ascription rule that constitutes the concept (the sense) expressed by the predicate '…is the inventor of the Maieutic' has the property of *being effectively applicable* to only one human being, namely, Socrates, reducing the domain of application to only one member. This is the same as attributing existence to the inventor of the Maieutic.[18]

Consider now the case of proper names. As we have seen, they should also have senses in the form of identification rules. Considering existence as the effective applicability of a possible semantic-conceptual rule in a chosen domain, the existence of the object referred to by a proper name should be established by the effective applicability of its possible identification rule, primarily in a proper contextualized domain of the external world.

Although this issue cannot be properly addressed without a deeper investigation of the nature of proper names, we can start by applying the Fregean-Russellian formal device to the foregoing view. In order to do this, we transform proper names into predicative expressions applied to

---

[17] It is easy to see that singular statements implicitly attribute existence to their objects, since a predicative statement with the form Fa could be written as $\exists x\,[\mathrm{Fx} \,\&\, (\mathrm{y})\,(\mathrm{Fy} \rightarrow (\mathrm{y} = \mathrm{x})) \,\&\, (x = \mathrm{a})]$ in order to make this attribution more explicit.

[18] Socrates lived in Greece from 470 to 399 BC. But usually the time and place of existence are abstracted when we talk about existence, since existence is essentially only the effective applicability of the conceptual rule, not the time of its applicability.

only one particular, showing then that the senses of names themselves can be reduced to the conceptual senses of predicative terms. A first step in the attempt to arrive at this is to transform the proper name into a predicate. Thus, 'Socrates' in 'Socrates exists' can be transformed into a predicate in the sentence 'There is something that socratizes,' or '$\exists x$(x socratizes).'[19] Taken literally, this suggestion is not only linguistically deplorable, but also formally deficient, since it leaves open the possibility that there is more than one Socrates.

Nevertheless, I think that '$\exists x$(x socratizes)' points in the right direction by suggesting that the existence of a name's bearer may be asserted by means of the conceptual senses of predicative terms. After all, the verb 'to socratize' can be seen as a kind of abbreviation of the predicative conceptual expressions included in the descriptions supposedly summarized by the proper name 'Socrates.' This is a reasonable strategy, insofar as we take seriously the bundle theory of proper names that was already fully present in one way or another in the writings of Frege, Russell, and Wittgenstein, though it has been made more explicit by P. F. Strawson and particularly John Searle. According to this theory, the whole sense of a proper name is given by a cluster of definite descriptions. Having this in mind, we might suggest that the attribution of existence to Socrates in '$\exists x$(x socratizes)' could be seen as an abbreviation of a set of predicative expressions like:

> $\exists x$ {x is inventor of Maieutic, x is mentor of Plato... x is Xantippe's husband}.

Of course, this is still inadequate, since it not only demands that all predicates must be satisfied, but leaves open the possibility that these predicates could be applied to more than one object. However, this fault can easily be remedied by means of the formal device that allows us to establish a minimum of at least one effectively applicable definite description:

> $\exists x$ {x and no other person invented the Maieutic, or x and no other person was the mentor of Plato or… or x and no other person was the husband of Xanthippe}.

[19] It was W. V-O. Quine who suggested using the name Pegasus as a way to change a name into a predicate such as 'the thing that pegasizes' (1948/9: 27).

Symbolizing the predicates '…is the inventor of the Maieutic' as P1, '…is Plato's mentor' as P_2, and '…is the husband of Xanthippe' as Pn, the above sentence can still be symbolically formulated as follows:

$$\exists x\,[(P_1x \,\&\, (y_1)\,(P_1y_1 \rightarrow (y_1 = x)) \vee (P_2x \,\&\, (y_2)\,(P_2y_2 \rightarrow (y_2 = x)) \vee \ldots \vee (Pnx \,\&\, (y_n)\,(P_ny_n \rightarrow (y_n = x))]$$

Here the supposed meaning of a proper name is disjunctively translated into the conceptual-senses of predicative expressions such as P_1, P_2… P_n, which according to our analysis are nothing but ascription rules expressed by predicates that we expect to be really applicable to one and the same thing. So analyzed, the attribution of existence to the object referred to by a proper name is made by saying that its sense, its identification rule, definitely applies in the assumed context. As this rule for the identification of a name was here analyzed in terms of a disjunctive set of rules for the application of predicates that must be applied to the same individual, we can easily explain existence as follows: The existence of the bearer of a proper name is the same as the effective applicability of at least one conceptual rule of a predicative expression to precisely one individual.

Of course, here it could be objected that such a descriptivist attempt to explain the meaning of a proper name is doomed to failure. This must be so, not only because the applied formal device is limited, but also because it amounts to some version of the bundle theory of proper names with its well-known difficulties, already persuasively pointed out by Saul Kripke, Keith Donnellan, and others...

However, such a conclusion would be too hasty, and there are at least three reasons to oppose it. The first is that, contrary to a current bias, Kripke's and Donnellan's objections have not discredit the most comprehensively developed versions of descriptivist theories, and some criticism has already been answered with considerable success by John. R. Searle (1983, Ch. 9). A second reason is that Kripke's alternative solution, the causal-historical view, could never be developed beyond a rough sketch.[20] These first two points lead us to the conclusion that bundle theory hasn't yet been definitely refuted.[21] Indeed, perhaps it just needs a stronger defense.

---

[20] There are less successful attempts, like Michael Devitt's interesting book *Designation* (1981).

[21] David Braun and Marga Reimer, two renowned specialists, made a balanced comparison of descriptivist and causal-historical views in their respective articles for the *Stanford Encyclopedia of Philosophy*. The results were inconclusive.

15. The existence of objects and its identification rules

The third and really conclusive reason that I can oppose to the anti-descriptivist view is that the above presented formal analysis is still a crude simplification when seen from the viewpoint of the new version of the bundle theory of proper names I have exposed in the Appendix to Chapter I. This version has, as I believe I have demonstrated, a much greater explanatory power than any previous theory, answering in a more nuanced way the most diverse counter-examples.

Briefly repeating what I said there, my view is the following. The traditional bundle theory of proper names defended by Frege, Russell, Wittgenstein, P. F. Strawson, John Searle and others has a severe limitation that has been overlooked: the bundles have *no internal order*. The theory does not tell us which descriptions or combinations of descriptions are more or less important or even why some seem to be very important for the application of a name, while others are obviously irrelevant for it. Definite descriptions are nothing but expressions of rules that should help us to connect a proper name with its reference. I called them description-rules. Regarding all this, my question was whether we cannot find the general form of a rule that we all implicitly know, which if applied to any bundle of descriptions associated with a proper name enables us to recognize the most relevant ones and decide in what ways the satisfaction of these descriptions makes this proper name applicable to some referent.

When searching for the general form of a rule, the first thing to do is to classify the descriptions. There is a sensible, ordinary-language method to use in order to begin with: check how encyclopedias treat well-known proper names. We can thereby easily distinguish *fundamental* from merely *auxiliary* descriptions, which are accidental. In doing this we see that proper names are first and foremost attached to two fundamental forms of description, which I call *localizing* and *characterizing description-rules*. Here is how we can define them:

(A) *Localizing description-rule*: This is the description that gives the spatiotemporal *location* and *career* of the object referred to by the proper name.

(B) *Characterizing description-rule*: This is the description that gives the characteristics of the object that we consider the *most relevant* to be referred to by the proper name – which gives us the *reasons* to use the name.

Consider, for instance, the name 'Adolf Hitler.' Here is what is said about its bearer in the first paragraph of a *Wikipedia* article:

> Adolf Hitler (20 April 1889 – 30 April 1945) was born in Braunan an Inn, Austria. Later he was a German politician and leader of the Nazi Party. He was Chancellor of Germany from 1933 to 1945 and Führer of Nazi Germany from 1934 to 1945. As effective dictator of Nazi Germany, Hitler was at the center of World War II in Europe and the Holocaust.

It is usual in encyclopedias that the first thing we find is an abbreviation of the localizing description-rule, followed by an abbreviation of the characterizing description-rule, stating the reason why we remember the name. What follows in the *Wikipedia* article (as in many others) are more or less relevant details and explanations. We find a variety of definite and indefinite descriptions that are more or less irrelevant: accidental, auxiliary descriptions. Examples of them are that Hitler was 'the lover of Eva Braun,' 'the son of Alois Hitler and Klara Pölzi'[22], 'the person called "Adolf Hitler",'[23] 'the boy who was sent by his father Alois to the *Realschule* in Linz in September 1900.' All this information given by encyclopedias will also be found in a more extended form in biographies.

You find a similar pattern if you search in encyclopedias for other proper names like 'New York,' 'USA,' 'Eiffel Tower,' 'Niagara Falls,' or 'Milky Way.' Of course, there are also the proper names of ordinary persons who are not famous enough to mention in encyclopedias. But the basic mechanism of reference remains the same. It is not difficult to see that the relevant information is given by their localizing descriptions and by the usually much more scattered characterizing descriptions. So, in most cases, if you wish to know who Sam is, you can probably get relevant information from his identity card, drivers license, employment record, police record (if any), school reports, club records… and most of all from details given by him, by his family and friends about his personality, character, education, interests, abilities, relationships,

---

[22] In some cases, like 'Queen Elizabeth II,' the family and even genetic origin is part of the localizing description, although this isn't necessarily so (See Appendix to Chapter II).

[23] What symbolic form a proper name receives is contingent. What makes this form necessary is the identification rule that we attach to it. In a possible world where the name attached to the identification rule for the name Hitler was attached to the name Hartman, this different name would mean what we mean by the name Hitler.

accomplishments, faults, etc., which are linked together by just one spatiotemporal career.

Now, my suggestion is that, although a conjunction of the localizing and the characterizing descriptions isn't required in any possible world, as Kripke has clearly shown (1980: 62), an *inclusive disjunction* of the two fundamental description-rules must in some degree be satisfied to enable a proper name to refer to its object in any possible world. John Searle perceived this point many years ago when he wrote:

> …if none of the identifying descriptions believed to be true of some object proved to be true of some independently located object, then that object couldn't be identical to the bearer of the name. (1969: 169)

Indeed, if we discovered records of a man named Adolf Hitler who was born in Gusental and lived in Austria from 1634 to 1689, worked as a shoemaker and had no political interests, we could safely conclude that he wasn't our Adolf since he does not satisfy any of the disjunction.

Moreover, two other complementary conditions should be added. First, a *condition of sufficiency* that must be satisfied: the disjunction of these two fundamental descriptions must be at least sufficiently satisfied in order to enable a proper name to refer to its object in any possible world. So, you can imagine a possible world where there was no World War II but where Adolf Hitler was born on 20 April 1889 in Braunan an Inn as the son of Alois and Klara Hitler. However, he had the same career as Adolf Hitler up to the point where he was not rejected but rather accepted by the Vienna Academy of Fine Arts in 1907, becoming a rich landscape painter who lived a long uneventful life. In this case, we are inclined to say that this person is *our* Hitler in this counter-factual situation, although he satisfies only the localizing description-rule, and even this only partially. But he already satisfies the inclusive disjunction sufficiently.

The second important condition is that of *predominance*, demanding that a possible bearer of a proper name should satisfy fundamental descriptions in a more complete manner than any other competitor in a possible world, since by definition the bearer of a proper name cannot be more than just one specified object. Thus, suppose that in a very similar possible world there were twins Adolf and Rudolf Hitler, both born on… 20 April 1889… but only Rudolf went to Berlin, served in World War I and later headed the Nazi Party, starting World War II and the Holocaust, while Adolf became a farmer in his native Austria. We would choose Rudolf as the true Hitler, despite his different name, since Rudolf satisfies the disjunction of conditions belonging to the identification rule for our Adolf in a much stronger way than the name of his twin brother presented

by the auxiliary description 'the person called "Adolf Hitler".' This shows once more the low relevance of auxiliary descriptions.

Finally, it is important to add that the object of reference belongs to the nearest relevant class 'C' that does not mix with the contents made explicit in the localizing condition (here, not being a politician, but being a human being).

Bringing all this together, we are able to propose the following *general form* of any identification rule for proper names, a form that must be satisfied by any bundle of descriptions associated by the linguistic community with a given proper name:

> *General form of the identification rule for proper names:*
> A proper name called 'N' has a bearer
> *iff*
> it is something that belongs to the nearest relevant class of referents C, so that *more than any other entity of the kind* C it *sufficiently satisfies* at least the conditions set by:
> (A) its localizing description-rule,
> and/or
> (B) its characterizing description-rule.
> (Auxiliary descriptions can be added as helpful symptoms for the identification[24]).

Now we can apply this form to any well-known bundle of descriptions that we associate with a proper name in order to have its *identification rule.* When we link the general rule form with the bundle of descriptions associated with the proper name 'Adolf Hitler', we get the following identification rule for this person:

> The proper name 'Adolf Hitler' has a bearer
> *iff*
> the bearer is something that belongs to the class of human beings, so that *sufficiently* and *more than any other* human being he satisfies the following inclusive disjunction of conditions:
> (A)being born on 20 April 1889 in Braunan an Inn… living the last part of his life in Germany… dying on 30 April 1945 in Berlin, and/or

[24] Remembering that there is no sharp boundary between fundamental and auxiliary descriptions.

(B) being the leader of the Nazi Party… dictator of Nazi Germany from 1934 to 1945… the person most responsible for World War II and the Holocaust.
(He would very probably also satisfy helpful auxiliary descriptions like being 'the lover of Eva Braun,' 'the person called "Adolf Hitler",' etc.)

This summarized identification rule gives us the *core meaning* of the proper name 'Adolf Hitler.' If we try to imagine an Adolf Hitler who does not minimally satisfy the fundamental localizing and/or characterizing conditions, we see that this is impossible. This was the case of the Adolf Hitler born in Gusental in 1634, who was a peaceful shoemaker and had nothing to do with politics. Surely, he could not be the person in a political socio-historical context whom we always mean by the name 'Adolf Hitler,' but someone else with the same name.

This example also outlines the lack of relevance of auxiliary descriptions. Suppose that the Adolf Hitler born in Gusental in 1634 satisfies many of the best-known auxiliary descriptions: he was the lover of an Eva Braun, he was the son of an Alois Hitler and a Klara Pölzi, the person called 'Adolf Hitler,' the boy who was sent by his father Alois to the *Realschule* in Linz… The feeling elicited by these strange discoveries would be of deep puzzlement, not persuasion. For his Eva Braun could not be the well-known Eva Braun who also committed suicide in the *Bunker*… and even that his parents had the same name as those of the infamous Adolf Hitler would be merely a remarkable coincidence… (He could not, it is true, satisfy the description 'the author of *Mein Kampf*'; however, more than an auxiliary description, this is already part of the full characterizing description of our Adolf Hitler.) Anyway, at no point will this change our belief that he is *not* the person we are trying to identify.

Since so understood the identification rule simply defines which object among all others owns the proper name by establishing the definitional criteria for identifying the proper name's bearer in any possible world, it unavoidably also *applies in any possible world where the name's bearer exists*, satisfying the fundamental requirement of the Kripkean definition of a *rigid designator* (1980: 48). The individually taken definite descriptions belonging to the bundle, particularly the auxiliary ones, on the other hand, being only *loosely* associated with the identification rule, can refer to other objects in different possible worlds and are therefore only accidental or flaccid designators.[25]

[25] One could object that rules are changeable and that if we change the identification rule, it ceases to be a rigid designator, unaware that auxiliary

Moreover, one can insert a name correctly in a sufficiently vague discourse without knowing more than auxiliary and indefinite descriptions, even when they are wrong, as Kripke realized. This is the case at least insofar as these descriptions are *convergent* (rightly classified), making in this way what we should call a *parasitical reference*, which can be helpful in several ways. For instance, if someone already knows that Hitler was 'some dictator' or erroneously thinks that he was 'a military general,' this person already classifies him correctly as a man of power can already apply the name correctly in sufficiently vague contexts and possibly be corrected and learn more about him.

Now, the existence of an object referred to by a proper name is the effective applicability of what can be called the *identification rule* of the proper name in its (in most cases) proper contextual domain. We know that Hitler existed because we know that his identification rule was effectively applied, hence applicable, in the political-historical context of Europe in the first half of the twentieth century. Moreover, what allows us to say that the bearer of the proper name 'Hitler' exists is that the property-tropes that belong to this object *satisfy an identification rule that by this reason has the property of being effectively applicable to it, a property that is actual if the rule is instantiated in some mind, but that would be only potential if the rule were never instantiated in any mind* (what is almost impossible to imagine in the present case, but would be easily imaginable concerning an object like a primitive animal living in a distant planet.) This property of the potentially existing identification rule is a higher-order property of the object, endowing it with existence in the real world and not as something only conceivable.

16. Existence of spatiotemporal locations: indexicals

Finally, there is the problem of the application of the proposed analysis of existence to the reference of those singular terms that change their reference according to the context: the indexicals. I will consider them only very briefly. Take simple statements with indexicals as (pointing)

---

descriptions can be changed as much as one will. Nonetheless, if we change the fundamental rules so that the set of possible worlds to which the proper name applies can be distinguished as a different one, we are not applying the same proper name anymore. However, you may introduce changes like additions to the fundamental description-rules insofar as this only specifies the identification better, and thus affecting nothing essential, only adding the application or non-application to possible worlds where the applicability of the rule was in an earlier stage *indeterminate*. (*Cf.* Appendix to Chapter I, sec. 7)

'There is a raven,' *'Here* is cold,' 'It rained *yesterday*,' *'I am* tired,' *'I am here now*'... The indexicals minimal task is to indicate some spatiotemporal location relative to the speaker. Thus, 'here' points to the place where the speaker is, 'now' to the moment when the speaker speaks, 'yesterday' to a period of time, the day before the day of the speaker's utterance… And regarding indexicals like 'I,' 'she,' 'he,' 'they,' there is more to say than just this. Surely, these personal pronouns have more semantic content than just a plain spatiotemporal location, but this does not matter to us now.

Consider now the indexical statement *'There* is a raven,' said when one found only one raven there. How should we analyze it? Of course, we can transform 'There' in the definite description 'the spatiotemporal location pointed to (or contextually shown) by the speaker when he utters the word,' which expresses a one-foot localizing identification rule followed by the countable predicate, the sortal '…is a raven' with its ascription rule. But in order to show our existential commitment, we need more. We need to analyze the definite description replacing the indicated spatiotemporal location by the predicate '…is in time t and place p' symbolized by L, replacing then the predicate '…is a raven' with the symbol R. With help of this we can symbolize 'There is a raven' as $\exists x\ [(Lx\ \&\ Rx)\ \&\ (y)\ ((Ly\ \&\ Ry) \rightarrow y = x)]$, which means: 'There is precisely one x that is in L and is an R.' Although the location L figures here as a predicate, the condition of unity (any $y = x$) makes it a singularized spatiotemporal location supposedly also analyzable in terms of tropes (See Appendix to Chapter III, sec. 3).

There is another common way to expose our existential commitment in indexical statements. It is when we add to them a sortal predicate, as in 'that raven there' in the sentence 'That raven there is flying' or 'this chair' in the sentence 'This chair is comfortable.' In these cases, we consider the phrases 'that raven there' and 'this chair' as referring to only one specific object, distinguishing it from all others. Hence, these phrases work *as singular terms* and must be analyzed as expressing identification rules. Replacing '…is a raven there' with R and '…is flying' with F, we can also formalize it as the existential statement $\exists x\ [Rx\ \&\ (y)\ (Ry \rightarrow y = x)\ \&\ Fx]$.

Indexical statements are important because when we use them the language, so to speak, 'touches' the world, which makes indexicals the indispensable roots of reference. Because of this, although the sense still determines its reference, we can find here a double direction of fit. First, with the help of our sensory cognitions, we create the identification rule for the indexical that is for the first time used in a determinate context. Once formed, this identification rule (a Fregean sense) determines the

spatiotemporal location, often together with the kind of object characterized by the sortal. Now, this new identification rule can be so established that it can be reapplied (not only later, but immediately thereafter), soon forgotten, or maybe interpersonally conventionalized by association with a non-indexical singular term of our language, normally a definite description. To this description, others can be later joined, building that bundle of descriptions able to flexibilize the referential work to many diverse circumstances which is typically abbreviated as a proper name.

17. Advantages of the higher-order view of existence

There are several advantages in conceiving existence as a higher-order property, that is, as a higher-order trope. One is that it gives a straightforward answer to what seems odd in the traditional forms of the ontological proof of God's existence. So, according to Descartes, once we accept the definition of God (1) as the being with all perfections, and that (2) existence is a perfection, we must conclude (3) that God exists (1978, V: 65). But if existence is a (tropical) meta-property of objects and not a proper intrinsic first-order tropical-property constitutive of them, differing in this way from perfections like infinite goodness, omniscience, and omnipresence, which should be intrinsic properties of God, the ontological proof is doomed to failure (*Cf*. Frege 1874, sec. 53).

However, the greatest advantage of conceiving existence as a higher-order property is that we will not have problems with the *denial* of existence. Suppose that existence were a first-order property of an object. In a sentence like 'Vulcan does not exist,' the negation of existence should then be applied to the object itself, and we would first have to identify the object in order to deny that it has the property, the trope of existence. That is, if in order to identify an object, we first had to admit that it exists, we would be caught in a contradiction: we would have to admit the existence of Vulcan in order to deny its existence.

According to our Fregean view, this contradictory conclusion isn't necessary, because all we do by denying the existence of Vulcan is to assert that the ascription rule that forms the concept of Vulcan doesn't have the meta-property of being effectively applicable in its proper contextualized domain of physical objects. Only to illustrate the point, we could analyze the sentence 'Vulcan does not exist' as a shorthand way of saying:

> ~∃*x* [(*x* is a small planet orbiting the Sun between Mercury and the Sun) & (y) (if *y* is a small planet orbiting the Sun between Mercury and the Sun, then *y* = *x*)].[26]

What belongs to the scope of '~∃*x*' are concepts constitutive of the identification rule, which in this illustration consists of an ascription rule for a predicate that can be applied to only one and the same object. What '~∃*x*' does is just to deny that this identification rule has the property of being effectively applicable to the corresponding physical object, which is to deny that an object existing only in our minds has the (meta-)property of also existing in reality.

18. Ubiquity of existence

The understanding of existence as the effective applicability of (semantic-cognitive) conceptual rules allows us to explain the almost unlimited extensions in the application of this concept. Why given that existence is primarily attributed to properties and objects of the outside world or of psychological states, are we also allowed to say that supposed entities like hypothetical and fictional ones exist? Some believe that even contradictory objects exist. We can even say that everything exists, including all that can be conceived – at least as something that can be conceived. And even of existence itself, it can be said that it exists. Indeed, it seems that in one way or another *everything exists*. How can this be possible?

Concerning supposed entities, we need to distinguish at least two kinds: *hypothetical* entities that experience hasn't yet shown to exist or has shown not to exist, and *imaginary* (including fictional) entities. Beginning with the first group, it is clear that we can find a sense in which they exist. Although the planet Vulcan has been shown not to exist in the real external world, its most proper domain, it surely has existed in the domain of the minds of many astronomers in the past who searched for it, as a hypothetical object… and it still exists in our minds, as a merely imaginary object.

For Frege, this would be a problem. But this is no problem for our proposed view because our identification rules can also have the existence-endowing property of being applicable, at least partially, in imagination, that is, only in the dependent domain of conceivable things that we consider as possible or even plausible candidates for existence in the external world. If I imagine the hypothetical planet Vulcan orbiting the

[26] This is again a didactic simplification (See Appendix of Chapter I).

Sun, I apply the identification rule for that proper name (even if in a vague, sketchy, deficient way) to a merely conceivable state of affairs. Indeed, the French astronomer Le Verrier, who first named the planet, even had a precise identification rule according to which Vulcan should be a small planet orbiting close to the Sun at a distance of 21 million km, which he mathematically calculated in order to explain by means of Newtonian laws the perihelion precession of Mercury's orbit. He applied this rule in the domain of what is conceivable, which means that Vulcan 'existed' in the restricted domain of the imagination of Le Verrier and other astronomers in his time, though not in its most proper domain – that of a concrete object, a planet belonging to the external world.

Consider now the case of purely fictional entities. Ivan is a character in Dostoyevsky's philosophical novel *The Brothers Karamazov.* He never existed in the real world; but he can be said to exist in the fictional world created in this novel, which is from the start fictional. In this domain, Ivan is the son of Fyodor Pavlovich and has two brothers, Dimitri and Alyosha. Ivan is a cerebral as much as a weak character, taking refuge from the inevitable confrontations of life in contemplation and inaction and creating resentful justifications for this; in the end, under the weight of his own conflicts, he descends into madness. These and other elements form parts of the rule for Ivan's identification. We say that he exists in the story, insofar as this rule is effectively applicable only to him within this proper fictional domain. Differing from the case of hypotheses, existence in a fictional world excludes from the start existence in the real world. That Ivan said to Alyosha: 'let the worms devour one another' is true in its fictional domain, as this statement is really made in the novel. But this utterance has no existence in the domain of the real external world, where it would be a displaced truth-bearer since the novel was not written to fit into it.

Saul Kripke gave examples of cases of fictional-fictional characters like Gonzago (2013: 250), who is a personage in Shakespeare's *Hamlet* as a fictional character created by Hamlet in his play within a play 'The Murder of Gonzago.' There is a hierarchy here. We may say that Gonzago exists in a third-order domain of Hamlet's play, requiring the effective applicability of a proper identification rule in this same domain. This third-order domain is supported by the existence of the plot of the fictional play *Hamlet,* forming a second-order domain. This play is in turn supported by the identification of some writer and writings in the first-order domain of our self-sustaining fundamental real empirical world.

As with other merely imaginary entities like winged horses and unicorns, existence is here affirmed within a domain that is *dependent,*

derivative or *extended* (Kripke 2013: 81), being supported by the fundamental form of existence, which concerns the effective applicability of cognitive rules in the domain of the *real* external (physical) or internal (psychological) world. Existence in these forms of usage is *parasitic* to the fundamental sense, though retaining its basic features (also Searle 1969: 78-9). In traditional philosophy, it was common to use the word 'being' instead of 'existence' for merely conceived existence. But I suspect that the real intention was often to underline the importance of conceived entities, underplaying or obstructing its derivative, parasitic character.

What about the attribution of existence to contradictory imaginative conceptions like that of a round square? This case seems really too hard to accept. We cannot combine the rule of identification of the square with the rule of identification of a circle so that both can identify one and the same thing, since they are from the start incompatible. We cannot do this even in our imagination. Because of this impossibility, we must recognize that in a literal sense a round square cannot reasonably exist: we cannot have a contradictory combination of conceptual rules, because it cannot form a possibly applicable rule combination. Since conceptual ascription rules are what constitute their cognitive meanings, this conclusion agrees with our strongest intuition: contradictions do not exist because they lack cognitive meaning.[27]

Finally, what about existence? Can we say that existence itself exists? Surely, we know that existence exists in the sense that we know that the concept-word 'existence' is effectively applicable to the property of effective applicability of conceptual rules in the most diverse domains, telling us that this property of effective applicability exists. This means that existence exists in the sense that we can build a meta-meta-rule of existence, whose criterion of application is the effective applicability of our metaconceptual rules made for the attribution of existence as the property of effective applicability of lower-order conceptual rules. Since there are meta-conceptual rules of existence which are effectively applicable (since entities belonging to their varied domains exist), the

[27] However, if the assertion that there are round squares were merely an equivocal manner of saying that we can syntactically combine the adjectives 'square' and 'round,' that is, a misleading way of saying that there is a syntactical rule allowing the combination of these incompatible words, then it makes some sense to attribute existence. But in this case, what we are trying to say will be more correctly expressed by the meta-linguistic sentence: 'The rule for constructing the phrase "round square" is applicable, and therefore, the phrase "round square" exists as a grammatical construction.' The Meinongian *Sosein* is reduced here to the recognition of a syntactical triviality.

meta-meta-rule – which demands the effective applicability of meta-rules attributing effective applicability to first-order conceptual rules – is also effectively applicable. Consequently, it is safe to conclude that existence itself exists. Well, then, does the existence of existence also exist? Surely: since the meta-meta-rule of existence is effectively applicable to meta-rules of existence by saying that the latter are effectively applicable to the first order conceptual rules, insofar as the latter ones are effectively applicable, we can conclude that a meta-meta-meta-rule of existence (affirming the existence of existence in itself) is also effectively applicable to the meta-meta-rule of existence, making the latter consequently existent. Of course, one can continue acknowledging the existence of the existence of existence and so on, in an infinite regress, which is virtuous since it can always be stopped.

19. Answering some final objections

According to many present theorists, existence is a first-order predicate. A statement like 'Horses exist' should be analyzed in a form similar to 'Horses are animals.' Since they have developed objections against the traditional second-order view, I will answer at least some of them, as they were formulated by Collin McGinn (2000b: 21-30). The answers can be helpful in clarifying my own standpoint.

The first one is against Russell's proposal that to say something exists is to say that a propositional function – a property, a concept – is true for at least one instance. Roughly stated, the objection is that for one object to *instantiate* a property this object must already exist, an admission that would make Russell's view circular, since it must already presuppose the existence of objects instantiating the property. For instance, if 'Mars is a planet' is true, it presupposes the existence of the planet Mars to instantiate the property expressed by '…is a planet' in order to make the sentence true. Summarizing, there must already be existent objects in order to instantiate the properties ascribed to them by our conceptual words.

This objection works insofar as one holds a Kripkean view of objects bearing proper names, since for him they cannot be defined by their own properties (1980: 52). Once we have analyzed an object as a widely accessible cluster of tropes displaying compresence, the objection appears to us in a different form. Since not only the ascriptive rules of predicative expressions, but also the identification rules of nominal terms are for us conceptual rules, our position should be generally stated as saying that existence is the effective applicability of any semantic-cognitive rules in some chosen domain or context. However, since these rules also apply to

objects as compresent clusters of tropes, this means we cannot conceive any object as being given – that is, as existing – without *simultaneously* conceiving its identification rule as effectively applicable to it. Thus, for instance, the existence of a concrete object like the planet Mars is nothing but the effective applicability of its identification rule in its proper astronomical context. This means that we cannot separate the existence of the object in its proper context from the effective applicability of its identification rule in the same context, since this is what warrants the object's existence. Now, if we assume that the attribution of truth to a singular predicative statement results from the applicability of the identification rule added to the applicability of the ascription rule, the attribution of properties and the admission of the object's existence are *conceptually correlative and cognitively simultaneous*. Moreover, as the truth follows from the combined application of the first two rules, it is wrong to insist that the attribution of truth requires the attribution of any property prior to the attribution of existence to the trope-property and the object as a cluster of trope-properties. The conclusion is that the flaw in McGinn's objection lies in the assumption that we can separate the instantiation of a property by an object from the attribution of existence to this same object.

Now to the second of McGinn's objections: uninstantiated properties are said to exist. But in order to exist, an uninstantiated property must fall under a higher-order property attributing its existence. This higher-order property must also exist, which means it must fall under a still higher-order property and so on infinitely. Consequently, the attribution of existence as a higher-order property is impossible, because it requires an infinite regress of properties to allow the attribution of existence.

My answer is that I agree (partially) with the diagnosis, but not with the prognosis. The effective applicability of a semantic-cognitive (conceptual) rule in its most proper domain not only endows its reference with existence, but is in itself a second-order property or trope that can also be said to exist. And furthermore, a semantic-cognitive rule that is only imaginatively applicable not only endows its reference with existence in an imaginary domain, but can also be said to exist. The trope-property of existence exists, which means that we can say that the second-order property of effective applicability of a conceptual rule can be the object of a third-order rule predicating its effective applicability, and so on indefinitely. This, of course, leads to an infinite regress. However, it is a *virtuous* infinite regress, since the applicability of a conceptual rule such as existence is already warranted by the application to it of a higher-order rule, and we don't need to bother with all the unlimited further

applicabilities of applicabilities or existences of existences that the first existence-endowing rule can generate. The mark of a virtuous regress is that we may *stop it without loss* when we feel that we do not need further steps to what we intend to explain, and this is the case here (See Appendix of Chapter III, sec. 2).

The third objection is that there are statements ascribing existence to particulars, such as 'Venus exists,' that resist the traditional paraphrase. We have already answered this objection in our treatment of proper names as conceptual identification rules.

But there are other objections. Consider the statement 'Something exists.' Although this is a true statement, McGinn believes that it is not paraphrasable in terms of the higher-order view, since there is no property to be instantiated here, and if we try to translate into the standard form we get the gibberish '$\exists x$(…x).'

The answer to this objection is too easy. What 'Something exists' means is that there is at least one trope or tropical construction out of tropes that exists without a further determination on our side. That is, we can say that there is some semantic-conceptual rule that is applicable to some domain of entities, even if this rule remains *unspecified*. This possibility is even shown by our logical symbolism on an elementary level, since we can symbolize an undetermined property such as, say, Φ. In this way we can translate 'Something exists' symbolically as $\exists x(\Phi x)$. But there is nothing wrong with $\exists x(\Phi x)$. Paralleling existential universalization, we can reach this result by considering singular existential statements like 'Venus exists.' So, calling Venus V, if it is true that '$\exists x(Vx)$' this implies that some property exists or '$\exists x(\Phi x)$, namely, that *some* conceptual rule is effectively applicable. This assumption of cognitively undetermined properties is harmless.

McGuinn reminds us that there are also more complicated statements that seem to resist a higher-order understanding of existence, like:

1. Some cities are purely imaginary.
2. Some of the things you are talking about do not exist.
3. There are things that do not exist…

Nonetheless, we can easily explain the predication of existence in such statements, insofar as we do not confuse the domains of application of the semantic-cognitive rules involved.

Thus, statement (1) means that some cities that exist in the imaginary domain exist only in this domain. Hence, the effective applicability of rules allowing us to identify the imaginary cities of Chloe and Valdrada in

the contextual domain of the book *The Invisible Cities* is sufficient for the attribution of existence to them in that purely fictional context. Statement (2) means that some things you are talking about exist only in imaginary domains, but not in the external world, that is, there are identification rules that are effectively applicable only in the unreal domain of one's own discourse. For instance, the identification rule of the name 'Vulcan' in the statement 'Vulcan is red' is only applicable in the speaker's (or hearer's) imagination. Finally, statement (3) means that there is at least one thing that exists only in the mind but not in external reality. Indeed, it seems obvious that the identification rule for some objects and therefore for at least one of them, though effectively applicable in an imaginary, only conceivable domain, isn't effectively applicable in the domain of external reality.

The last of McGinn's objections is that according to the higher-order view, nothing can exist without falling under some property or other, which rules out the existence of a thing that has no properties – a 'bare existent.' However, our empiricist commitment makes us see this not as a weakness, but rather as a further anti-metaphysical advantage of our understanding of the higher-order view.

20. Reference again: a metaphysical excurse (Mill)

It is instructive to consider what happens when we compare the famous phenomenalist view of J. S. Mill, according to which 'matter' or 'substance' is nothing but 'permanent possibilities of sensation' with our view of existence in terms of the effective applicability of conceptual rules. The results will be no less speculative than Mill's phenomenalism, but they may be telling.

Mill's great epistemological question was: If all that is experientially given to us are sensory phenomena, how can we justify our belief in the existence of an external world, an objective world constituted by substance or matter? – An external world that can exist even when there is no observer at all to perceive it?

Mill's answer to the question was a development of Berkeley's unofficial view, according to which things that we know to exist when we are not perceiving must be nothing more than things that we are certain we would perceive under suitable circumstances.[28] As Berkeley wrote:

[28] In accord with Berkeley's official view, things that are not actually perceived by us exist because they are continuously being perceived by God. (Urmson 1983)

> Existere is percipi or percipere… The horse is in the stable, the books are in the study as before. (1707-8, *Notebok A*, 429)

> The table I write on, I say, exists, that is, I see and feel it; and if I were out of my study I should say it existed – meaning thereby that if I was in my study I might perceive it, or that some other spirit actually does perceive it. (1710, I, sec 3)

According to this view, *esse* is not only *percipi*, but also *percipi possi*. In a more explicit manner, what Mill suggests is that:

> Matter or substance is not made up of actual sensations, but of groups of permanent (or guaranteed or certified) possibilities of sensation.

Mill justifies his identification of matter or substance with permanent possibilities of sensation in the following way. First, these possibilities of sensation are *conditional certainties*: they are not mere epistemic possibilities, but firm conditional expectations that are in direct or indirect ways based on experience. They are *permanent* in the sense that, once suitable circumstances are given, they would always be experienced insofar as they are said to exist. And they are *guaranteed* or *certified* in the sense that we have good reasons – observational or not – to have a firm expectation that under suitable circumstances they will be experienced again and again. This does not mean that the groups of permanent possibilities of sensations would depend for their existence on our past experience of them, because if that were so, they could not exist without us as subjects of knowledge, and we would fall like Berkeley into some radical form of idealism (Berkeley 1710, 1713). This was not Mill's intention. As he explains:

> We mean [by permanent possibilities of sensation]… something which exists when we are not thinking of it; which existed before we have ever thought of it, and would exist if we were annihilated; and further that things exist that we never saw, touched or otherwise perceived, and things which never have been perceived by man. (1979, X: 178-177)

Thus, it is clear that Mill wished to avoid idealism: the permanent possibilities of sensations would exist even if cognitive beings able to perceive them never existed.

These permanent possibilities are for Mill *objective*, differing from our actual constantly changing sensations, which are subjective. They are objective because they are grounded, he thinks, in our common public world, which makes us able to interpersonally agree on their existence. For

him, even if different persons cannot have access to the same sensations, they can have access to *the same* possibilities of sensation. As he writes:

> The permanent possibilities are *common* to us and to our fellow creatures, the actual sensations are not… The world of possible sensations succeeding one another according to laws is as much in other beings as it is in me; it has therefore an existence outside me; it is an external world. (1979, X: 181-2, my italics)

This is in summary Mill's view on the nature of matter – a view that always seemed to me as much deeply suggestive as contentious.

I think there is a serious confusion in Mill's view, which can be made clear when we compare his insights with those of Berkeley. According to the non-official Berkeleyan view, the external world is constituted by *sensations* whose experience is continually (permanently) possible for us, even if we are not there to experience them. But if this is so, the material objects constituting the external world cannot be reduced to simple 'groups of permanent *possibilities* of sensation,' for *possibilities* as such, permanent or not, cannot be qualitatively distinguished one from the other in the same way as one material object can be distinguished from another. *Material objects can be qualitatively very different from each other, they are multiple and varied, while possibilities are always the same, namely, mere possibilities*. Consequently, possibilities (of sensations), permanent or not, cannot be the same as material things. Keeping this in mind, the only feasible way to express the Berkeleyan insight in Mill's terminology seems to me to use it in the characterization of material objects, as follows:

> *Material objects* (or substances) are nothing but multiple and varied *groups of sensations* whose effective experience is permanently (or guaranteed or certified to be) possible.

This would meet the requirement of multiplicity and diversity proper to material objects and their presentations because each material object would be constituted by *innumerable groups of sensations* that under suitable circumstances could always be possibly distinctly experienced. But if the *permanent possibility* of sensations is not the material object, what is?

I believe it is a way to point to the external *existence* of the material object. This answer emerges when we consider Mill's view in the light of my reconstruction of Frege's concept of existence, according to which existence is the *effective applicability* of a conceptual or semantic-cognitive rule. If this is so, it seems that the permanent (guaranteed,

certified) possibility of groups of sensations could be approximated to the *existence* of such groups of sensations and the last ones to material objects; these warranted groups of sensations would be the same as the criterial configurations warranting the applicability of the rule. Consider the expressions:

1. Permanent (guaranteed, certified) possibilities of groups of sensations.
2. Effective experienceability of groups of sensations.

Expressions (1) and (2) say the same thing in different words. Now, compare them to the following expressions of existence in our reconstruction of Frege's view:

3. Effective applicability of a conceptual rule.
4. Effective applicability of a conceptual rule to groups of given sensory-perceptual contents.
5. Effective applicability of a conceptual rule to given (independent) criterial configurations or tropes.

Although (4) is only a variation of (3), it seems clear that when we interpret existence as (4) we are saying something at least equivalent to (2): the effective experienceability of groups of sensations. Since (2) is only a different way to say (1), the permanent (guaranteed, certified) possibility can be approximated to existence. One could even suggest:

> *Existence* is the *effective* (permanent, guaranteed, certified) *possibility* of groups of sensations.

The point in question is made clearer when we consider the general structure of our conceptual rules of ascription and identification. We already know that these rules have the form of semantic-criterial rules that bring us to some usually pre-reflexively achieved semantic cognition, given by the satisfaction of variable subjective criterial configurations (supposedly) by means of their match with objective criterial configurations, which should be nothing but configurations of external tropes. Now, when we interpret these variable criterial configurations as being the same as Mill's groups of sensations, as we have reconstructed them, we can speak of existence as the effective, guaranteed, certified, permanent possibilities of groups of sensations as consistent with the effective applicability of a conceptual rule. Here an example can be

helpful: In order to be applied to a real located object, the conceptual rule for the concept *chair* demands the satisfaction of criterial configurations. These criterial configurations are established by the definition of a chair as a seat with a backrest made for only one person to sit on at a time, which we could decompose in terms of subjective sensory criterial configurations that must be satisfied by matching objective criterial configurations or configurations of given external tropes. But the criterial configurations (the dependent ones, at least) could be reduced to groups of sensations whose experience is permanently (guaranteed, certified as) possible.

Now, Mill's insights can help us deepen our reconstruction of the Fregean concept of existence. A material object exists only:

(i) when its conceptual rule is effectively applicable, but this effective applicability is only the case when
(ii) criteria for the application of its identification rule can be objectively given to us at least in the form of groups of what we may call *independent, external contents of sensation* whose experience is warranted or permanently possible. Moreover, as Mill also suggested,
(iii) this possibility of experience must be (at least in principle and indirectly) interpersonally accessible by allowing agreement in the description of the experience;
(iv) this experience can be more or less direct;
(v) it is (usually) independent of our will; and
(vi) it is also experienced as following causal laws regarded as typical of things belonging to the external world.

It seems that all these things together contribute to building the condition of an effective application of a semantic-cognitive rule in the domain of the external world – they are contributing to warrant the attribution of *external existence*.

There is, however, an important and seemingly fatal objection to Mill's view of matter, which is made more serious by the Berkeleyan correction I made above.[29] It is that the group of sensations or configurations of sensory criteria that satisfy a conceptual rule are by their nature inevitably *psychological*. It seems clear that even sensations or contents of sensations that are warranted as permanently possible must be psychological in a dispositional way. This means that if we follow this path, we end up

[29] I believe that Mill's confusion in the definition of matter was in fact an attempt to evade the objection of idealism open to Berkeley.

falling into some form of Berkeleyan idealism in which there is no objective, external material world to be contrasted with our subjective world of sensations or sensory criteria. No really independent non-mental external trope needs to be there to match the apparently satisfied dependent criterial conditions, as suggested in statement (5). It is true that, as Mill noted, his possible sensations are independent of our will, that they follow the regularities of nature, even that they appear to be interpersonally accessible under circumstances that warrant their experience (under suitable circumstances they are described as being experienced simultaneously by different subjects, etc.). However, all these things do not seem to help because of the possibility of skeptical scenarios: they can all be unwittingly imagined, as in the dreams. They seem, therefore, insufficient to perform the magic of turning sensations *qua* sensations into something they aren't, namely, supposed elements of a non-mental objective external world of material objects with their own tropical-properties. This is an important objection, whose answer will be given only in the final chapter of this book, as a consequence of our discussion of the adequation theory of truth in its relation to direct realism.

Notwithstanding, I can now anticipate something of the way I intend to deal with the problem. Having in mind the suggested view of existence, we can ask: What warrants an object's external existence or reality? One answer could be: the joint satisfaction of conditions (i) to (vi) by (5) and nothing more. This would be all that we need to identify the external reality with the contents of our experiences, for there is no way to verify whether or not there is some radical skeptical truth concerning our whole external world, which under normal circumstances makes radical skeptical doubt senseless. (Ch. VI, sec. 30)

An associated question is: What is in this context an *external material object*? A too daring answer would be: the external object (as it is thought) must be the identification rule in itself, insofar as it is effectively applicable; in this way, the multiplicity and diversity of objects would be explained by the multiplicity and diversity of identification rules... However, this cannot be, since a semantic-cognitive rule is also something essentially mental, and we are definitely not what Plato called 'friends of ideas.'

Looking for a less daring answer, we can suggest that what we understand as the material object is not the semantic-cognitive rule, but is supposed to have the same structure as this rule *projected in a specular way onto the external world*. There is a reason for this suggestion: It seems that only something with a structure similar to its semantic-cognitive rule would be able to give unity to the multiple and variable criterial

configurations by means of which external entities are able to give themselves to us in our experience of them. Figuratively speaking, if the semantic-cognitive rule has the form of a tree with branches whose ramifications end in criterial conditions dependent to the rule, then the object of its application, as we believe it to be, must have the structure of an inverted specular tree with branches whose ramifications end in independent criterial configurations that (supposedly) should match the corresponding subjective criterial configurations. Furthermore, these objective criterial configurations should be nothing but external tropes and constructions out of them (objects, properties, facts). Of course, this objective structure should be putative, so that the rule could always be improved or corrected as a response to new information regarding such specular objective counterparts. (Ch. VI, sec. 34)

21. The reference of a sentence as its truth-value

Now we leave our speculative excurse and come back to the more tangible Fregean semantics, considering what he has to say about the reference of a sentence. Here I have no compliments to make. Frege was the author of the insane idea that the references of sentences are their truth-values, so that the thoughts expressed by them should be modes of presentation of truth-values.

How did he reach this strange conclusion? There are several reasons. First, he notes that sentences are independent, saturated, closed; they work in ways similar to those of names, and a truth-value is also closed, since it does not require complementation. Second, he says that the search for truth is what brings us from sense to reference. Third, he notes that sentences without reference lack truth-value: 'Vulcan is a warm planet' has no reference and for him no truth-value, since this hypothetical planet has been shown not to exist. Fourth, he also noted that conforming to the principle of compositionality – according to which the whole is a function of its parts – the reference must be what remains unchanged after we change the senses of a sentence's components without changing their references. This is what happens, for instance, if we replace 'Napoleon lost the Battle of Waterloo' with 'The man of destiny lost his last battle.' Since the references of the sentence-components do not change, the reference of the whole sentence likewise does not change. Moreover, the truth-value of both sentences remains the same: The Truth. Hence, their reference must be their truth-value. The conclusion of all this is that in extensional languages the references of sentences must be their truth-value (1892: 34). For Frege, all true sentences have only one reference, which is the abstract

object *The True* (*das Wahre*), while all false sentences also have only one reference, which is the abstract object *The False* (*das Falsche*).

However, there are a number of well-known embarrassing objections to Frege's identification of reference with truth-value that in my opinion completely disqualify his view. A first objection is that, contrary to any healthy intuition, Frege's proposal frontally contradicts the meaning we normally give to the word 'reference.' It is intuitively obvious that the sentence 'Napoleon was born on Corsica' refers to something very different from the sentence '2 + 2 = 4,' even if both are true. Moreover, if you replace 'Venus is a planet & the Earth is a planet' with 'Mars is a planet & the Earth is a planet,' both composite sentences remain true because of the truth of the partial sentences, but the reference of 'Venus' is totally different from the reference of 'Mars,' what runs against the principle of compositionality. Another objection is that we expect the references of components of our sentences to be on the same ontological level as the sentences' references. But for a Fregean, this could not be the case: the reference of the name 'Napoleon' is the Napoleon of flesh and blood, while the reference of the sentence 'Napoleon was born on Corsica' must be the abstract object called The True. Moreover, Frege's solution violates his own principle of compositionality. If the reference of a sentence is its truth-value, it cannot be established by its parts, since a truth-value has no parts. And even if it had parts, then all objects referred to by names in true sentences should be parts of The True, which would hardly make sense. There are also serious substitutability problems with Frege's explanation of the references of sentences. The first is that if all true sentences refer to The True, and the name 'The True' also refers to The True, then in the conditional sentence 'If it rains, then water falls from the sky,' we can replace 'it rains' with 'The True.' But the result will be the sentence 'If The True, then water falls from the sky,' which should be true but is in fact unintelligible (Black 1954: 235-6). A second and fatal problem of the kind is that a multitude of obviously false identities between true sentences should be true. For example, 'Paris is a city = snow is white' should be a *true* assertoric composite sentence, since the two sentences refer to the same thing: The True. Under critical scrutiny, Frege's view shows itself to be hopeless.

The most charitable interpretation is that Frege uses the word 'reference' as truth-value because it is what counts, because the word *Bedeutung* (meaning) in German, more than in English, also means *relevance*, pointing to *semantic relevance* or *meaningfulness* (*Cf.*

Tugendhat 1992b: 231).[30] Indeed, truth-value is of decisive relevance for logic, because it is what must be preserved in valid arguments. The logician does not need to know more than truth-value regarding the referring function of the sentences he is dealing with in order to evaluate inferential possibilities.

A main problem with this interpretation is that it contradicts expected principles of Frege's own theory. Since the reference (*Bedeutung*) of the parts of a singular sentence (general and singular terms) can be seen as their references in a literal sense (the concept and the object that can fall under it), truth-value as relevance satisfies the principle of compositionality in an odd, non-linear form, since relevance is normally only an adjective applied to truth-value. This is different from the principle of compositionality applied to senses in which the whole and its components are linearly arranged in the same semantic domain. The attempt to tell us that a reference is mere qualification attributable to it is equivocal and confusive.

Finally, when we take the truth-value for the reference of a sentence, this view can be – and in my judgment really has been – utterly misleading from an epistemological standpoint. Since truth considered as in some way belonging to thought has nothing to do with anything that can reasonably be understood as the reference of our statements, calling truth-value 'the reference' contributes to placing the relation between language and the world virtually beyond semantic reach.

22. Logical structure of facts

The Fregean account of the references of sentences as their truth-values turns out to be still less acceptable if we consider that a much more natural alternative is available, which, as Sir Anthony Kenny has noted, was not even mentioned by Frege (Kenny 2000: 133). This alternative, which the logical atomism of Wittgenstein and Russell tried to explore, consists in the appeal to *facts*. Since it is *prima facie* much more plausible that the references of sentences are facts, it is important for us to investigate the logical structure and ontological nature of facts.

Considering first the logical structure of facts, a plausible view is that they correspond to the logical structure of the thoughts representing them, assuming that these thoughts are what declarative sentences express when logically analyzed, at least in accordance with the context of the linguistic practices where they occur. Nevertheless, even respecting linguistic practices we can go further, considering that they are placed within the

[30] See Frege, Letter to Russell of 28.12.1912.

factual language in general and accepting a form of atomism in which the bottom line of the analysis is the exposure of the logical components of what is stated in singular sentences where we can find identification rules of singular termini associatively used with ascription rules of predicative expressions. Singular empirical statements such as 'Frege has a beard' and 'The cat is on the mat' belong to this bottom line and respectively represent *facts* that should have the logical structure depicted by *Fa* and *bRc*.

Elements *a*, *b* and *c*, as singular terms, refer to individuals constructed as clusters of appropriate compresent tropes, while *F* and *R* would also be seen as designating tropes, usually complex tropes forming complex criterial configurations dependent on the clusters to which they are tied. The ties between *b*, *R* and *c*, and between *F* and *a*, in turn, are only pseudo-relations, since admitting their existence as relational tropes would generate an inevitable infinite regress. As we already noted, individuals and their property-tropes are linked by 'non-relational ties' without any ontological addition (*Cf.* Appendix to Chapter III, sec. 1). Indeed, what could be the relational ties between the application of the ascription rule of '…was bearded' to Aristotle with the already applied identification rule of Aristotle in the fact represented by the statement 'Aristotle was bearded'?

We should also pay attention to the somewhat trivial rule of analysis according to which *we should not accept singular terms – and even candidates for this function – as components of complex predicative expressions* (I say candidates, intending sentences like 'The Minotaur has two horns'; since the horns are individuals, they must be referred to by singular terms in the completely analyzed sentence). (*Cf.* IV, sec. 7) Thus, for instance, in a sentence like 'Stockholm is the capital of Sweden' we should not view '…is the capital of Sweden' as a predicate, since Sweden is a proper name. Also inadequate would be to analyze 'the capital of Sweden' as a definite description contextually referring to Stockholm in our world, so that the analyzed sentence would have as its relational predicate '…is (the same as)…' The most appropriate analysis would be to consider '…is the capital of…' as a relational predicate completed by the proper names 'Stockholm' and 'Sweden,' separating the relational trope from the compresent bundles of tropes referred to by the proper names. Proper names are stronger identifiers than definite descriptions and should therefore be preferentially singled out in the logical analysis of thought.

Furthermore, it also seems possible to analyze proper names and definite descriptions using Russell's technique of transforming them into quantified predicative expressions, insofar as to a limited degree this device mirrors the neodescriptivist theory of proper names defended in this

book, a similar procedure being possible regarding general terms. Anyway, such sub-sentential terms normally do not need to be analyzed when our task is to analyze sentences, since they are the proper elements of sentences, except when they are not what they seem to be, as in the case of nominalizations.

Finally, we have composite facts represented by our extensional language, along with the general (universal, existential) facts to be analyzed as having the same structure of sets (conjunctions, disjunctions) of singular statements that make up general (universal, existential) statements, which, as we already noted, can be reduced to associations of singular predicative and relational statements. (I think that the philosophical problem of a hidden *lingua mentis* ends up in elements like those briefly pointed out in this section).

23. Ontological nature of facts

If we accept that the references of sentence-senses or thoughts are facts, then from an ontological perspective what empirical sentences represent must be empirical facts, most typically located in the external world, though possibly also located in the inner mental world. This assumption speaks for the correspondence or adequation theory of truth, according to which empirical facts are truth-makers normally seen as complex contingent arrangements of elements in the world, that is, *usually contingent tropical arrangements associating tropical individuals and property-tropes*.

However, this assumption conflicts with Frege's anti-correspondentialist view of truth. According to him, a fact would simply be a *true thought* (1918: 74). Following similar anti-correspondentialist lines, in a very influential article, P. F. Strawson suggested that empirical facts are mere 'pseudo-material correlates of the statement as a whole' and not something in the world (1950: 6). According to him, empirical facts, unlike *events* or *things*, are not spatiotemporally localizable ('the world is the totality of things, not of facts'). One reason for this is that the description of a fact usually begins with a *that-clause*. For instance, I can say 'the fact *that* the book is on the table,' but not 'the fact of a book on the table.' On the other hand, the description of an event typically lacks a that-clause: I can say 'the event of a tsunami in Japan,' but not properly 'the event *that* there was a Tsunami in Japan.' Facts are for Strawson what statements (when true) state, not what statements are about. They are

> not, like things or happenings on the surface of the globe, witnessed or heard or seen, broken or overturned, interrupted or prolonged, kicked, destroyed, mended or noisy. (1950: 6)

The same is for him the case with states of affairs and situations.[31] Finally, to give a striking example, the event of Caesar's crossing the Rubicon *occurred* in the year 47 BC, while the fact that he crossed the Rubicon did not occur in the year 47 BC, but it is still a fact today, since facts simply do not occur (Patzig 1980: 19-20).[32]

An easy way to dispose of this argument could be the following. We need a word to describe the condition in the world that makes our thoughts true. The word 'fact' is available. So, why don't we use it *stipulatively* in order to designate the truth-maker, whatever condition it is?[33]

However, it seems clear to me that even this stipulative way to circumvent the problem is avoidable, since it is not difficult to show that the problem exists only in the imagination of philosophers. To begin with, of course not everything we may call a 'fact' is empirical in the usual sense of the word. It is hard to assign empirical status to the fact that 2 + 2 = 4, even if its supposed non-empirical character can be an object of controversy.[34] And we can say 'It is a fact that the Sun is *not* green,' although this seems to me only a linguistically modified way to say 'There is *no* fact that the Sun is green' or 'The fact that the sun is green does not exist.' What I want to defend here is that there is a privileged sense of the word 'fact' that involves references to more or less obvious empirical facts, particularly so-called observational facts, which should be considered objectively real: they exist in the external world and can be seen as the ultimate truth-makers of their statements.

To begin with, it is good to remember that there is a well-known and very convincing reason to think that facts can be constituents of the empirical world. This is that many facts are said to act causally. Consider the following sentences:

(1) The fact that the match was scratched caused the flame.
(2) Thomas died because of the fact that he forgot to turn off the gas.

---

[31] Without offering a justification, Strawson writes: 'a situation or state of affairs is, roughly, a set of facts, not a set of things.' (1950: 8)

[32] For an important reply, see J. L. Austin, 'Unfair to Facts' (1961, Ch. 5). It seems to me at least curious that the posthumously published arguments of Austin against Strawson's view have had so little impact.

[33] John Searle once proposed something approaching this answer: '…we neither have nor need a thick metaphysical notion of "fact." Anything sufficient to make a statement true is a fact. Thus the fact that there are no three-headed cats is as much a fact as the fact that the cat is on the mat.' (1998: 392)

[34] See Appendix of Chapter III, sec. 4.

(3) Because of the fact that today is a holiday, the class will be canceled.
(4) The fact that Caesar crossed the Rubicon had important historical consequences.

It does not seem possible that pseudo-material correlates (which I suppose to be abstract contents) can be causally active in the empirical world, producing these effects. But conceding the empirical nature of facts (1) to (4) solves the problem in obvious ways. Scratching a match is a *fact-event* causing a flame. The *situational fact* created by Thomas' forgetting to turn off the gas caused his death. The *fact-circumstance* that today is a holiday causes the cancellation of a class. The *fact-event* of crossing the Rubicon established a *state of affairs* that causally determined decisive political changes in the Roman Empire.

Furthermore, I have a key-argument to regenerate the idea that empirical facts are correlates of true thoughts, as the classical correspondence theory of truth has held. According to the view I propose, empirical facts are contingent tropical arrangements in the external and/or internal world in general. Similar would be the case with facts apparently as simple as those referred to by sentences like 'Frege had a beard,' 'The Eiffel Tower is in Paris,' and also facts constituted by combinations of such facts.

My argument against Strawson's opposition between non-spatiotemporal facts and spatiotemporal events begins by showing that there is a serious confusion in his argument. He treats facts (as much as states of affairs and situations) as *opposed* to events. His schema is:

FACTS	*x*	EVENTS
Pseudo-material correlates		Spatiotemporal phenomena

But this can easily be contested. We begin to be suspicious when we perceive that every event can be called a fact, but not every fact can be called an event. For instance: I can replace 'the event of the sinking of the Titanic' with 'the fact of the sinking of the Titanic,' but I cannot replace 'the fact that Mt. Everest is more than 8,000 m. high' with 'the event of Mt. Everest being more than 8,000 m. high.' Strawson's opposition isn't symmetrical. Now, since events can be called facts, it is much more reasonable to consider events as particular kinds of facts than to oppose the two, as Strawson did. Indeed, *my proposal is that the word 'fact' is an* umbrella term *that encompasses events, occurrences, processes, as much*

as situations, circumstances, states of affairs, etc. And the reason for this proposal is that we can call all these things facts, but we cannot call all these things states of affairs or events. We see that events are *sub-types* of facts and that linguists could classify the word 'event' as a hyponym of the word 'fact.' Considering things in this way, we can distinguish two great sub-classes of facts:

1. STATIC FACTS: Can be formal or empirical, the latter when clearly located in space and time. As a whole, static facts *do not change while they last*. Typical of static facts is that the relationships between their tropical components do not decisively change during the period of their existence. They are truth-makers of a static kind. And ordinary language has names for them: they are called (with different semantic nuances) 'states,' 'situations,' 'conditions,' 'circumstances,' 'states of affairs,' 'ways things are,' etc.
2. DYNAMIC FACTS: These are always empirical. They *change while they last*. The relationships between the elements constitutive of them change decisively during the period of their existence, so that they have a beginning, followed by some kind of development that comes to an end after a certain amount of time. We will see that they work as truth-makers of a dynamic kind. And ordinarily they can be called (with different semantic nuances) 'events,' 'episodes,' 'occurrences,' 'occasions,' 'processes,' 'transformations,' etc.

Facts said to be formal, like the fact that $7 \times 8 = 56$, are static in the harmless sense that they do not need to be considered as spatiotemporally located. They are not of concern to us here. Many facts are empirical and static, insofar as the relationships between the elements constitutive of them do not change during their existence. Static facts are usually called 'states,' 'situations,' 'conditions,' 'circumstances,' 'states of affairs'… with different nuances of meaning. Examples of static facts are my *state* of poor health, the *situation* that I am lying in bed, the *circumstance* that the airport is closed, the *state of affairs* that the Mona Lisa is in the Louvre or that the Earth orbits the Sun. The Earth's movement of revolving around the Sun does not count because it is an internal cyclical relationship that remains the same during the fact's existence: as a whole, this state of affairs does not change while it lasts (although each orbital period counts as an event).

Dynamic facts, on the other hand, can be called 'events,' 'episodes,' 'occurrences,' 'occasions,' 'processes,'… They are defined by changes in their overall composition and in relations among their elements during the period of their existence. World War II, viewed a process, for instance, began with a rapid expansion of the territories dominated by Nazi Germany and was marked by events like the Battle of Britain, the Battle of Stalingrad and the Normandy invasion – it had an unforeseeable history. Dynamic facts are usually called events when their duration is comparatively short, occurrences when their duration isn't as short, processes when their duration is longer. Examples of events are an explosion or a lightning flash in a storm. An example of an occurrence is a volcanic eruption. The process of global warming is a very slow natural process, slower than the process of economic globalization. We can predict the stages of many events and processes, although many are also unpredictable in such a way that (unlike static facts) we cannot grasp them in their integrity before they end. Important is to see that all these things can be individually called events, occurrences, occasions, happenings, processes… and also *facts*, since they are all nothing but empirical facts – truth-makers of a dynamic kind.

We are now able to find what seems to be the real reason why we use a that-clause in the description of facts, but not in the description of events. When we speak of dynamic facts, we do not use a that-clause. Thus, we can speak about the event of Caesar's crossing the Rubicon, but not about the event *that* he crossed the Rubicon. We can speak about the process of climate change, but not about the process *that* the climate changes… But this isn't the case regarding static facts, which are typically (though not necessarily) described as beginning with that-clauses. So, I can speak about the state of affairs *that* my book is on the table or *that* I am lying on the bed, although I can also speak about the state of affairs of my book being on the table and of my lying on the bed. The conclusion is that if that-clauses have some function it is that of *excluding dynamic facts and emphasizing static facts*. Moreover, since the hyperonymic term 'fact' can be applied to both – static facts as much as dynamic facts – it is reasonable to suppose that this term *inherits* the property of being used indifferently, with or without a that-clause. Indeed, you can say, 'It is a fact *that* Mount Vesuvius is located near Naples' (referring to a state of affairs), as much as 'It is a fact *that* Mount Vesuvius has erupted' (referring to an event). And we can also say: 'Caesar crossing the Rubicon was an event', as much as 'It is a fact *that* Caesar crossed the Rubicon,' referring less precisely to the event. We can summarize these relationships in a schema:

(a) Static facts (states of affairs…): can be well stated with or without a that-clause.
(b) Dynamic facts (events…): cannot be well stated with a that-clause.
(c) Facts in general: admit both cases, because being all-embracing they do not differentiate between (a) and (b).

Now, what about the fact that Caesar crossed the Rubicon? Isn't this fact timeless? The answer is that this is a good case of a misleading statement. In most cases, it is not understood as the description of an event, but as an illustrative way of referring to a *static social fact*: the state of affairs established by the movement of Caesar's army onto Roman territory, violating the law that prohibited this and forcing the Roman state to declare war against him. Only occasionally is the phrase 'crossing the Rubicon' understood in its literal sense, as the physical event of crossing the river, which comprises Caesar's sequential locations in relation to the river from t_1 to t_n.

Due to the nature of dynamic facts like events and processes, we say that they not only *are*, but also *occur* in time, while of static facts we only say that they are *located* in time while they last. It seems, therefore, that because philosophers such as Strawson did not realize that events are subtypes of facts, seeing only that we may say of events that they occur in time, they hastily concluded that only events (and things) are located in time, opposing them to timeless facts. But that this isn't true can be shown even by inter-substitutivity *salva veritate*: it is correct to say that the event, the occurrence of Caesar's crossing the Rubicon, *was* a fact and that this fact occurred in 47 BC, as a concrete dynamic fact. On the other hand, the static social fact, the political state of affairs established by Caesar's crossing the river was far more enduring. Being a static fact, it was the political *situation* that led, as is well-known, to the fall of the Republic. However, it seems clear that the state of affairs brought about by the crossing of the Rubicon was spatially limited to the Roman Empire and temporally limited to the time from Caesar's crossing the Rubicon to his coronation as Caesar and up until his assassination. It was not something that existed in Greenland or that endured until the present, even if in a misleading way our ordinary language can be confusive by allowing us to use the present tense to speak about historical facts.

The relevant conclusion is that by having the broadest scope, the so often vilified word 'fact' remains the ideal candidate for the role of ultimate truth-maker in a correspondence theory of truth. Facts are universal truth-makers.

24. Church's slingshot argument

As already noted, for Frege a sentence's reference is its truth-value. To refute the charge that this view is implausible, the Fregean logician Alonzo Church devised a slingshot argument. He wanted to show that by means of inter-substitutability of co-referentials we can prove that the most diverse sentences can only have a truth-value as their reference.

Church's argument is equivocal, but telling. Its basic assumption is that when one constituent expression is replaced by another, so that their *partial references* (the references of their singular terms) are *interchangeable*, the reference of the whole sentence does not change. I will begin by explaining his slingshot argument, underlining its supposedly co-referential definite descriptions (Church 1956: 25):

1. Sir Walter Scott is the author of *Waverley*.
2. Sir Walter Scott is the man who wrote the twenty-nine *Waverley* novels altogether.
3. Twenty-nine is the number such that Sir Walter Scott is the man who wrote that many *Waverley* novels altogether.
4. Twenty-nine is the number of counties in Utah.

According to him, if it is plausible that sentences (2) and (3) are, if not synonymous, at least *co-referential* sentences, then (1) has the same reference as (4). Since (4) seems to concern a fact completely different from (1), it seems that the only thing left as the same reference is the truth of both sentences. Hence, The True is the only referent of all these sentences.

However, the argument proves to be unsustainable when we pay attention to what should be the real reference of each singular term of these sentences. In sentence (1) the proper name 'Sir Walter Scott' and the definite description 'the author of *Waverley*' are two singular terms expressing different modes of presentation of the same human being. These modes of presentation make what we could call two *partial* references to Walter Scott, namely, references that must be partial relatively to the whole reference of the sentence. In sentence (2) again, the nominal expression 'Sir Walter Scott' and the definite description 'the man who wrote the twenty-nine *Waverley* novels altogether' both refer in different ways, that is, partially, to the same Walter Scott. The third sentence is the tricky one. Its reference is unclear: Walter Scott? The number 29? Both in one? The combination Scott-29? The answer appears when we paraphrase sentence (3) so that it gives back in a transparent way

its complete informative content. Now, carefully considering the confusing sentence (3), we see that the only way to reveal its content in a transparent way without any addition or loss of sense is to split the sentence into the following conjunction of two sentences: (5) '29 is the number of *Waverley* novels *and* Sir Walter Scott is the man who wrote that many *Waverley* novels altogether.' Sentence (5) makes explicit all the content wrapped up in sentence (3). For the sake of clarity, replacing in (5) '=' for 'is (the same as)' and '&' for 'and,' we can still unpack (3) as:

6. (29 = the number of *Waverley* novels) & (Sir Walter Scott = the man who wrote the many *Waverley* novels altogether).[35]

That is: Sentence (3) confusingly compresses nothing less than a conjunction of two identity sentences, each with its own proper partial references given by the singular terms flanking their identity signs. They are the number 29 in the first sentence and *Walter Scott* in the second. Finally, we come to the analysis of sentence (4): '29 is the number of counties in Utah,' which means the same as the identity sentence (7) '29 = the number of counties in Utah.' Here, each singular term that flanks the identity sign has the number 29 as its partial reference. So analyzed, the derivation appears as:

1. Sir Walter Scott = the author of Waverley.
2. Sir Walter Scott = the man who wrote the 29 Waverley novels altogether.
3. (5) (29 = the number of Waverley novels) & (Sir Walter Scott = the man who wrote the many Waverley novels altogether).
4. (6) 29 = the number of counties in Utah.

Now, although all these sentences are true, Church's argument has by now lost its initial plausibility. Sentences (1) and (2) have as the partial references made by their singular terms Walter Scott under different guises. However, sentence (3) is a conjunction of two identity sentences, each with its own very distinct partial references. The object referred to by the flanking terms of the first identity sentence of (3) is the number 29 (as the number of Waverley novels), while the object referred to by the flanking terms of the second identity sentence of (3) is Sir Walter Scott (as the man

[35] This also gives back the whole sense of Church's still more convoluted original sentence: 'The number such that Sir Walter Scott is the man who wrote that many Waverley Novels altogether is twenty-nine.'

who wrote the Waverley novels). Finally, sentence (4) has as partial references made by its singular terms only the number 29 (as the number of counties in Utah), without referring to Walter Scott, as it should. That is:

> In the composed sentence (3), the second sentence of the conjunction is the only one that preserves as the partial reference made by its singular terms the references of (1) and (2), while (4) is an identity sentence that has as partial references made by its singular terms only the same partial references of the first sentence of (3). However, this is precisely what should not occur, because the preserved partial references have nothing to do with the partial references made by the singular terms of sentences (1) and (2) and the object referred to by them. Consequently, the whole references of these sentences and sentence (4) must be different.

In other words, we can say that in a surreptitious way the replacements slide equivocally from having partial references to Walter Scott in (1) and (2), to a Walter Scott, together with the number 29 in (3), and to the number 29 in (4). This means, according to the principle of compositionality applied to complete sentences, that the references of sentences (1) and (4) should indeed be very different. Initially, the flaw is not easy to spot, because sentence (3) contains both objects of partial references conjoined in a grammatically confusing way. We have the impression that the partial references of (3) seem to be something like an amalgam of Walter Scott and 29, say, a 'Scott-29,' while they are and must, in fact, be totally distinct. The replacements would only respect the compositionality principle, warranting the sameness of the sentences' references, if the argument could prove that the partial references of *all* the sentences could be replaced without furtively inviting the reader to conjoin in sentence (3) partial references to completely different objects.

25. Sub-facts and grounding facts

If we take the whole reference of the sentence as not a truth-value but a *fact,* we get much more intuitive results. In what follows, I will consider Church's intended derivation, not only to introduce facts as referents of sentences, but also to introduce a very useful distinction between *sub-facts* and *grounding facts*. As will be seen, this distinction fills a gap in Frege's explanation.

We need to distinguish at least two facts referred to by identity sentences. The first is the sub-fact: it is the *perspectival fact as the appearance immediately revealed through a particular mode of presentation expressed by the statement*. I will call it a *sub-fact* and make the diversified sub-facts the objective correlates responsible for differences in the modes of presentation constitutive of the different sentences' senses (thoughts, rules) concerning one and the same object, e.g., Walter Scott and the author of *Waverley*. This is why Church's sentences (1) and (2) can be seen as expressing different senses or thoughts. They evoke different perspectival sub-facts. They indirectly represent different sub-facts, since (i) *being* Sir Walter Scott is not the same thing as (ii) *being* the author of *Waverley* and (iii) *being* the man who wrote the 29 *Waverley* novels altogether… In this way, sentences (1) and (2) respectively show two different sub-facts that contain perspectival objects of reference that as such differ from one another. Using the term 'being' to indicate that we are speaking about a matching correlate, the *sub-facts* represented by:

(1) Sir Walter Scott is the author of *Waverley*.
(2) Sir Walter Scott is the man who wrote the 29 *Waverley* novels altogether.

Can be respectively represented as follows:

(1a) Being Sir Walter Scott ≠ being the author of the *Waverley* novels.
(2a) Being Sir Walter Scott ≠ being the man who wrote the 29 *Waverley* novels altogether.

These sub-facts are of contingent *differences* since Sir Walter Scott could have not written the *Waverley* novels or any novel in the first case, and he could have written a different number of *Waverley* novels in the second. (If you accept that there are relational tropes of identity, you should accept that there are here relational tropes of difference.)

Nonetheless, it is also clear that (1) and (2) are identity sentences. This is so because these sentences can be understood as referring under different guises to only one object, the person called Walter Scott, justifying the employment of the 'is' of identity. In this sense, sentences (1) and (2) represent an identity, which can be expressed simply by 'Walter Scott = Walter Scott.' That is, they can represent the self-identity of Walter Scott considered in full, as the ultimate bearer of all descriptions (under all possible perspectives) that we might intend to use to refer to it. Among the descriptions we associate with the name 'Walter Scott' we can

select 'the person with the title of Sir named "Walter Scott"' (that is, 'Sir Walter Scott'), 'the author of *Waverley*' and, certainly, 'the man who wrote the 29 *Waverley* novels altogether,' that is, the constituent expressions of (1) and (2). Now, this primary fact that Walter Scott is (the same as) Walter Scott (considered in full) is what I call a *grounding fact*. Characteristic of the grounding fact is that it must be able to unify all the sub-facts, all the facets revealed by its multiple modes of presentation. This is what remitting us to sentences of the form a = a make sentences with the form a = b identity sentences.

Now, consider one of these definite descriptions more carefully, for instance, 'the author of *Waverley*.' As we saw, the mode of presentation is intentional and internal, considering that the reference can be absent. But when the mode of presentation isn't empty, as in this case, it also exposes something external, evoking what I could spell out as '*being* the author of the *Waverley* novels.' This should be seen as an objective phenomenal entity, a *sub-object* mediating our reference to the object Walter Scott that belongs to the grounding fact of Walter Scott's self identity.

As well, 'the author of *Ivanhoe*' (who was also Walter Scott) is a mode of presentation of the sub-object '*being* the author of *Ivanhoe,*' though it ultimately refers to Walter Scott. Now, take the sentence:

(a) The author of the *Waverley* novels is the author of *Ivanhoe*.

This sentence evokes two different sub-objects that together form the contrastive sub-fact that being the author of *Waverley* is not the same as being the author of *Ivanhoe*. But this sub-fact also consists of two modes by which *the same* object is given, whose identity is the grounding fact that can be directly represented by the sentence 'Being Walter Scott [in full] = Being Walter Scott [in full],' where 'in full' here means that we are intending to consider all the conceivable modes of presentation of the object Walter Scott, far beyond the limited knowledge of this or that particular speaker.

Moreover, it seems clear that the sentence (a) must also be able to express the two thoughts representing the two kinds of facts considered. First, we have a *derived thought* expressible by the sentence (a_1) 'Being the author of Waverley novels isn't the same as being the author of Ivanhoe,' representing directly the sub-fact and indirectly the grounding fact. Second, we have the *basal thought* directly expressible by the

sentence (a_2) 'Being Walter Scott [in full] = being Walter Scott [in full],' representing the grounding fact directly.[36]

According to the foregoing analysis, when I say 'The author of *Waverley* novels is the author of *Ivanhoe*,' I am saying two things. First, by means of intentional modes of presentation, I am expressing the *derived thought* evoking a factual objective difference. This thought can be expressed by the sentence 'Being the author of *Waverley* ≠ (isn't) being the author of *Ivanhoe*,' representing a derived fact. Indeed, it is an objective factual difference that a person writing *Waverley* is not the same as a person writing *Ivanhoe*, even if they are both the same person (he was writing different stories at different places and times…). However, since when I say 'The author of *Waverley* is the author of *Ivanhoe*' I use an 'is' of identity, I also mean the *basal thought* expressible by the sentence 'The author of *Waverley* = the author of *Ivanhoe*,' indicating that under different guises I am presenting the grounding fact that 'Being Walter Scott = Being Walter Scott.' It is because of the two – the grounding fact along with the sub-fact – that identities of the kind a = b are able to express identities in their differences.

Now, assuming the kind of neo-descriptivism proposed in Appendix I of this book, we can make explicit the above-mentioned doubling of the presented facts by stating each of the four sentences of Church's reasoning as follows:

(1a) *Sentence expressing the derived thought representing the sub-fact*: Being Sir Walter Scott ≠ being the author of *Waverley*.
(1b) *Sentence expressing the basal thought representing the grounding fact*: Being Walter Scott [in full] = being Walter Scott [in full].

(2a) *Sentence expressing the derived thought representing the sub-fact*: Being Sir Walter Scott ≠ being the man who wrote the 29 *Waverley* novels altogether.
(2b) *Sentence expressing the basal thought representing the grounding fact*: Being Walter Scott [in full] = being Walter Scott [in full].

[36] I think that the mode of presentation of the sub-fact can be approximated with what defenders of two-dimensionalism call a *primary intention* (here called *derived* thought) while the mode of presentation of the grounding fact can be approximated with what they call a *secondary intention* (here called *basal* thought) (*Cf.* Chalmers 2002). Anyway, the present suggestion is clearly more perspicuous and natural.

(3a) *Sentence expressing the derived thought representing the sub-fact*: (Being 29 ≠ being the number of *Waverley* novels) & (Being Sir Walter Scott ≠ being the man who wrote the 29 *Waverley* novels altogether).
3b) *Sentence expressing the basal thought representing the grounding fact*: (Being 29 = being 29) & (Being Walter Scott = being Walter Scott).

(4a) *Sentence expressing the derived thought representing the sub-fact*: Being 29 ≠ being the number of counties in Utah.
(4b) *Sentence expressing the basal thought representing the grounding fact*: Being 29 = being 29.

The sub-facts show why the semantic contribution of each referential component in identities with the form a = b, due to the semantic-cognitive rules constitutive of the derived thought, can be different. The sub-fact that being Sir Walter Scott isn't the same as being someone who wrote 29 *Waverley* novels discriminates more than the sub-fact that being Scott isn't the same as someone writing the *Waverley* novels. And regarding true sentences, this discrimination isn't just a mentally considered mode of presentation, a cognitive rule, but also the representation of something objectively or factually given in the external world (corresponding to different 'ways the *object* gives itself to us,' using Frege's words). The above presented evocations of sub-facts all lead us to two grounding facts of identity showing how many different senses referring immediately to qualitatively different sub-facts refer mediately to something numerically identical. On the other hand, in sentences with the form a = a, such as 'the morning star = the morning star,' the sub-fact is already the identity 'Being the morning star = being the morning star.' The corresponding grounding-fact, additionally, may also be the same identity, if not the identity 'Being Venus = being Venus,' depending on the speaker's intention.

26. Taking seriously the sentence's reference as a fact

I think I have shown that the most plausible option concerning the nature of reference is to side with philosophers like Russell and the earlier Wittgenstein. These philosophers assumed that the reference of a statement is a *fact* – a fact that in the usual case is understood as a contingent arrangement of cognitively-independent tropical components commonly given (completely or partially) in the external world, although

they can also belong to an internal (psychologically accessible) reality. Facts would satisfy the Fregean condition that the reference of a sentence is an object: they are in some sense independent, complete, closed. They would satisfy the condition that thoughts expressed by sentences should also be *modes of presentation* of their references, the latter – particularly as sub-facts – being as numerous and diverse as their thoughts. Finally, unlike truth-values, facts would smoothly satisfy the principle of compositionality: they would always vary in accordance with variations in the references of the senses of component parts of the sentences as we understand them.

If we assume the answer given above, we are able to solve a vexing problem concerning which fact the thought expressed by a sentence refers to. Consider the following sentences:

1. The morning star is the morning star.
2. The morning star is the evening star.
3. Venus is the morning star.
4. Venus is the second planet orbiting the Sun.
5. Venus is the brightest planet visible in the sky.
6. Venus is the only planet in our solar system shrouded by an opaque layer of highly reflective sulphuric acid clouds.
7. The morning star is the only planet in our solar system shrouded by an opaque layer of highly reflective sulphuric acid clouds…

On the one hand, it is intuitively correct to say that each of these sentences refers to a different fact. Sentence (1) is tautological, proclaiming the factual self-identity of the morning star, while sentences (2) to (7) provide information on different factual contents regarding the planet Venus. On the other hand, since all singular terms composing these identity sentences have the same ultimate reference, the planet Venus, it also seems clear that in the end all these identity sentences must have the same reference, representing the same fact in the world. How can we reconcile these two seemingly correct views?

The answer departs from the distinction already made in the last section: first, there must be a privileged *grounding fact* able to be described that can be identified as the ultimate truth-maker of all these identity sentences about the planet Venus. Second, this grounding fact must in some way contain the facts immediately indicated by the different cognitive values of sentences (1) to (7) above as its perspectival *sub-facts.* My suggestion is that this last task can be accomplished by the references of identity sentences, insofar as the identification rules of their singular

terms are considered *in full*, including all their fundamental and auxiliary descriptions.

Now, assuming our proposed view of proper names' meanings as abbreviations of bundles of descriptions centered in those constituting their fundamental identification rules, then the proper name 'Venus' in full includes in its most complete content all the already known modes of presentation. This means that definite descriptions such as 'the morning star,' 'the second planet orbiting the Sun,' 'the brightest planet visible in the sky,' etc. can have their application made at least probable by applying the concept of Venus in full. (I say 'made at least probable' because, in the case of most identification rules, any particular description-rule of the bundle might be wrong and remain unsatisfied.) If this view is correct, then there is only one sentence that could describe the grounding fact as the ultimate truth-maker or verifier of any identity sentence concerning the planet Venus, including the sentences from (1) to (7) above. We can present it as the grounding fact (8) that being Venus with all its known sub-factual identificational inferences is being Venus with all its known sub-factual identificational inferences, represented by the basal thought expressed by the sentence:

(9) Venus [in full] = Venus [in full]

My contention is that rightly understood this sentence summarizes the most complete basal thought able to represent the single *grounding fact*, which considered in its entirety can be regarded as the truth-maker for any identity sentence about the planet Venus. (To represent sub-facts we have the already named derived thoughts.)

It is not hard to explain why things are so. If the full meaning of the proper name 'Venus' is understood as an abbreviation of the whole bundle of descriptions regarded as uniquely identifying its object (*Cf.* Appendix of Chapter I, sec. 4), then this proper name should include descriptions like 'the morning star,' 'the evening star,' 'the second planet orbiting the Sun,' 'the most brilliant planet visible in the sky,' 'the only planet in our solar system shrouded by an opaque layer of highly reflective sulphuric acid clouds,' and many others. Consequently, from the sentence 'Venus [in full] = Venus [in full]' we can inferentially derive sentence (2) 'The morning star = the evening star.' We do this simply by replacing the first occurrence of the name 'Venus' with the definite description 'the morning star,' which the name 'Venus' (in full) abbreviates, and the second occurrence of the name 'Venus' (in full) with the description 'the evening star,' which the name Venus also abbreviates. In a similar way, we can

obviously (inductively, at least) infer all the other above presented co-referential identities from (1) to (7). Thus, rightly understood the sentence 'Venus [in full] = Venus [in full]' should express the basal thought able to represent a fact complex enough to comprehend all the sub-facts represented by each of the thoughts expressed by the above sentences, which may be seen here as *contingent a posteriori*. (To convince yourself of this, look at the meaning of 'Venus' as presented in any encyclopedia, since it aims to offer an abbreviation of Venus in full.)

In order to better support what I am suggesting, I can also use numerical identities like the following:

1. $2 + 2 = 2 + 2$
2. $2 + 2 = 1 + 1 + 1 + 1$
3. $2 + 2 = 4$
4. $4 = \sqrt{16}$
5. $2 + 2 = (14 - 6) / 2$

Of course, here the identity sentence expressing the basal thought representing the grounding fact would be:

6. The number 4 [in full] = the number 4 [in full]

But could the sub-facts expressed by sentences (1) to (5) be derived from (6)? Obviously, the answer must be in the affirmative, since we are dealing with a deductive system. After all, I wrote the five sentences above simply based on deductive inferences from my knowledge of the grounding fact that being the number 4 = being the number 4!

However, one could still object that a sentence like 'Venus [in full] = Venus [in full]' is a tautology: a necessary truth. How could a necessary truth ground contingent truths like, 'Venus is the brightest planet visible in the sky'?

My answer is that for an idealized privileged user of the word (or an astronomer) who is supposed to know all the relevant information about Venus, this proper name expresses an identification rule that can be approximatively summarized as follows:

> IR-Venus: Our proper name 'Venus' has a bearer, *iff* this bearer belongs to the class of celestial bodies that satisfy sufficiently and more than any other the condition of being the second planet orbiting the Sun between Mercury and the Earth. (To this it is helpful to add very probably applicable auxiliary descriptions like 'the brightest

planet visible in the sky,' 'a planet somewhat smaller than the earth,' 'the morning star,' 'the evening star,' etc.)

As in the case of the Venus called 'Hesperus' (Appendix of Chapter I, sec. 10 (iii)), this is a kind of 'one-foot' identification rule, since the localizing rule is the only fundamental one and includes what would count in the characterizing rule (being a planet). For suppose we have as a characterizing rule 'a bright planet somewhat smaller than the earth.' In this case, one can imagine that if there were only one bright planet somewhat smaller than the Earth, this planet would be Venus, since one term of the inclusive disjunction of a fundamental identifying rule is already satisfied. But if this were true, since we can imagine a possible world where there is just one bright planet somewhat smaller than the Earth with an orbit outside the Earth's and no second planet, this planet should then be Venus, which is absurd. And as noted, the localizing rule contains the essential characterizing content: Venus as a planet. If Venus were to lose its atmosphere or a major share of its mass (or in a different possible world never had them), insofar as it had been discovered to be the second planet from the Sun and the Earth the third, it would still be our Venus! Indeed, so understood it seems that the identification rule for Venus is applicable in any possible world where the planet Venus can be said to exist or to have existed.

The case of Venus is somewhat like the case of the lines 'a⌒b-a⌒c' drawn to localize the center of a triangle without any call for a characterizing property; the characterizing description can be irrelevant or non-existent. By the same token, without the localizing condition established by the identification rule of Venus as the second planet, it would be impossible to identify Venus. The application of many other descriptions does not produce criteria, but only symptoms of the planet's existence, since they make the applicability of the descriptions only more or less probable. Auxiliary descriptions like 'the brightest planet in the sky' are symptoms, like 'the highly reflective clouds of sulfuric acid' that cause this brightness. If Venus lost its reflective atmosphere, it might cease to be the brightest planet, but would still not cease to be Venus. If Venus lost half of its mass but remained in the same orbit, it still would not cease to be Venus. But if for some reason Venus lost nearly all its mass and became a small orbiting object only a few miles in diameter, no longer large enough to be called a planet, we could only say that it once was Venus. If in a possible world Mercury *never* existed, Venus would be the first planet of the solar system and even if it were called 'Venus,' it seems clear that it would not really be our Venus, unless it had once been

the second planet from the Sun (Venus) for at least some period of time. Indeed, if in another possible world the second planet were hurled out of the solar system thousands of years ago (Kripke 1980: 57-58), it could still rightly be recognized as our Venus, since it once satisfied its identification rule. We see that the condition of sufficiency applied to the one-foot identification rule of Venus is more demanding than in the usual two-foot case. And we see that limits can be set even in a swampy terrain where vagueness prevails.

What I said about identity sentences also applies to other singular predicative and relational sentences. Consider the following ones:

1. Bucephalus was a material thing.
2. Bucephalus was a living being.
3. Bucephalus was a horse.
4. Bucephalus was a black horse of the best Thessalonian strain.
5. Bucephalus was a massive black horse of the best Thessalonian strain, owned by Alexander the Great.
6. Bucephalus: (355 BC – 326 BC) was the most famous horse of Antiquity; it was a massive black horse of the best Thessalonian strain, owned by Alexander the Great.
7. Bucephalus once swam across the river Granicus.

One could say that each of the first six sentences expresses different derived thoughts representing different sub-facts by means of increasingly detailed modes of presentation expressed by their respective predicative expressions. However, relative to them there is a grounding fact that in a summarized form is represented by the basal thought expressed by sentence (6), since the truth of all the others can be implied by the truth of this thought. Indeed, (6) is nothing but an abbreviated expression of the identification rule for Bucephalus, with a localizing and a characterizing description and by these means furnishing a summarized definitional criterion. The sub-facts represented by sentences (1) to (5) are all included in the grounding fact represented by sentence (6). These facts are the immediate satisfiers of the diverse modes of presentation of Bucephalus given by each sentence. And the progression from (1) to (6) increases the complexity, insofar as new relevant predications are added. Statement (7) ‘Bucephalus once swam across the river Granicus’ is a different case: the very contingent auxiliary description ‘the horse Bucephalus who once swam across the river Granicus’ isn’t a relevant part of the fundamental description-rule (even if he didn’t swim across the river, he would still be our Bucephalus).

Nevertheless, it can still be derived from (6) considered in full, since this is believed (by privileged speakers) to be historically the case.

27. The riddle of identity in difference

There is a final point concerning the relationship between the sub-fact and the grounding fact. It concerns the unsatisfactory way that Frege solved the puzzle of identity. As he wrote, unlike sentences with the form $a = a$, a sentence with the form $a = b$ is *informative* because it refers to the same object by means of different modes of presentation, by means of the different senses of *a* and *b* (1892: 26). However, we can still ask how this identity is possible, since the modes of presentation are different and since we are not intending to speak about the mere self-identity of the reference, as Frege also acknowledged. I call this 'the riddle of identity in difference.'

To see the problem clearly, consider again Frege's sentence (i) 'The morning star = (is) the evening star.' A more fully unpacked cognitive sense of (i) can be presented as:

> The brightest star in the *morning* sky, understood as referring to the second planet orbiting the Sun between Earth and Mercury (Venus) = (is) the brightest star in the *evening* sky, understood as referring to the second planet orbiting the Sun between Earth and Mercury (Venus).

Here I have not underlined non-definitional expressions of what I call *immediate senses* presenting perceptual sub-objects like the morning and the evening star, though I have underlined expressions of what I call *mediated senses*, which here are definitional. The immediate senses build the derived thought representing a sub-fact (that being the morning star isn't being the evening star), while the mediated senses essentially build the basal thought representing the grounding fact (that Venus is Venus). Here we have the hidden reason for the riddle of identity in difference: the immediate senses of the expressions flanking the identity sign in (i) are obviously different, but they both evoke the underlined mediated, in fact primary or leading sense (essentially building the basal thought that Venus is Venus) with the form $a = a$.

Obviously, this last sense, the basic thought that the second planet orbiting the Sun… is the second planet orbiting the Sun… is not yet the reference, since it is constituted by the expression of the self-identity of the cognitive identification rule constituting the core definitional sense of the name 'Venus' and its conventionalized surroundings (Venus in full). It is only because both expressions flanking the identity sign in (i) implicitly evoke the same proper identification rule for the planet Venus that we are

allowed to place an identity sign between them! In order to make the point still clearer we can appeal to the following schema:

Sentence:	The morning star	= (is)	the evening star.
Derived:	IR: the brightest	≠	IR: the brightest
thought:	star in the morning		star in the evening.
sub-fact:	Being the morning star isn't being the evening star.		

Basal	IR: The second planet…	=	IR: the second planet…
thought	(Venus)		(Venus).
grounding fact:	Being Venus is the same as being Venus.		

In sum: the singular terms 'Morning Star' and 'Evening Star' are responsible for the difference present in what I call the *immediate senses* of the descriptions (the Fregean senses) constituting a derived thought evoking a relational sub-fact showing the differences between two sub-references. Expressing the derived thought we describe the sub-fact as: 'being the brightest star seen in the morning sky differs in place and time from being the brightest star seen in the evening sky' (one can even point to the two opposite sides of the sky in which alternately one or the other appears every twelve hours). Furthermore, the 'is' understood as 'is the same as' is the only indication of the identity of the implicitly intended *mediated senses* building the basal thought expressed by the sentence 'The second planet orbiting the Sun between Earth and Mercury (Venus) = the second planet orbiting the Sun between Earth and Mercury (Venus).' These mediated senses have multiple guises that are implicit in the names flanking the identity sign in the statement 'Venus [in full] = Venus [in full]' expressing the basal thought that could be known in full only by specialists or idealized speakers. The statement expressing the derived thought is *contingent a posteriori,* while the statement expressing the basal thought can be seen as a *necessary priori*.

A somewhat different example is the sentence 'The morning star is Venus.' Here the schema is:

Sentence:	The morning star	= (is)	Venus.
Derived:	IR: the brightest	≠	IR: the second
Thought:	star at dawn		planet.
sub-fact:	Being the morning star isn't being Venus.		
Basal	IR: the second planet	=	IR: the second planet
Thought:	(Venus)		(Venus).
grounding fact:	Being Venus is the same as being Venus.		

It is by now clear that the identity expressed by sentences of the kind a = b is an identity in difference. This means that in fact we have two levels of sense or thought. The first is the derived thought. It represents the perspectival sub-fact with its sub-objects expressing a *difference* (Being the morning star isn't the same as being the second planet from the Sun). The second, intermediated by the first one and indicated by the 'is' of identity, is the basal thought representing the ultimate grounding fact that being Venus is the same as being Venus, which has the sub-facts as facets, as manifestations. The derived thought is contingent a posteriori, while the basic thought is a *necessary a priori* expression of a conditioned rule.

Now, how should we deal with cases in which the elements of the basal thought responsible for the identity, like the planet called 'Venus' in the statement above, lack a proper name? Consider the identities (i) 'Everest = Chomolungma,' (ii) 'a⌒b = a⌒c' (concerning Frege's example of two different ways to name the center of a triangle), (iii) 'Afla = Ateb' (the two names that Frege gave for the same imaginary mountain). In order to get an answer, we need to first consider that the derived Fregean senses are thoughts of a difference, evoking different contingent sub-objects. But these sentences also implicitly evoke a basal conjoining sense, a conjoining identification rule, which refers to what we might call respectively the 'Everest-Chomolungma,' the 'a⌒b-a⌒c,' and the 'Afla-Ateb,' which in fact are three new nominative expressions. The law of identity makes it obvious that:

(1) 'Mt. Everest is Chomolungma' so understood can be replaced by 'Everest[-Chomolungma] = [Everest-]Chomolungma,'
(2) 'a⌒b = a⌒c' can be replaced by 'a⌒b[-a⌒c] = '[a⌒b-]a⌒c,' and
(3) 'Afla = Ateb' can be replaced by 'Afla[-Ateb] = [Afla-]Ateb.'

These three replacing basal thoughts respectively represent the three different grounding facts as the full self-identities that they are. This is respectively what sustains the identities expressed by the 'is' in the sentences (i), (ii) and (iii).

We can apply a similar analysis to identities between concept-words of the form (*x*) (F*x* = G*x*). Consider the identity 'Heat in gases is molecular kinetic energy.' Note that the word 'heat' is ambiguous. It can mean a mere subjective feeling (heat1), like the feeling of increased bodily heat after exercise, which cannot be identified with molecular kinetic energy. But in the present case 'heat' means external temperature as it is normally felt by people (heat2). A third sense is independent of our sensations: it is heat as 'measured temperature' determined by thermometers (heat3) (in

the sense of heat2, our bodies serve as coarse, imprecise thermometers). Moreover, since molecules can have different masses and speeds, the most precise identity sentence would be 'Temperature in a gas (heat3) is the average kinetic energy of its molecules.' This sentence expresses two different modes of presentation of the same thing, that is, a derived thought that can be expressed by means of the following difference:

(i) Temperature in a gas (heat3) ≠ average kinetic energy of its molecules.

This secondary thought refers only to the sub-fact that the (macro-physical) temperature that we can measure with a thermometer (and feel as heat2) is something phenomenally different from the (microphysical) average kinetic energy of the molecules of a gas such as the air around us.

In a next step, we are able to consider the basal thought establishing a tautological identity based on conventions. This thought can be expressed by the whole more complete assertoric sentence:

(ii) [Average kinetic energy-temperature-] heat3 of a quantity of gas = average kinetic energy-temperature [-heat3 of a quantity of gas].

Now, we can read the sentence 'Heat in gases is molecular kinetic energy' as something made explicit by sentence (ii), which can be read in two ways: (a) considering only what is outside the brackets as *explicitly emphasized*, which expresses the derived thought of a difference and represents the sub-fact of the difference above; (b) *emphasizing the whole, including what is in brackets*. Understanding (ii) as (b), what we have is a basal thought referring to a grounding fact of definitional self-identity. This identity requires as an assumption the acceptance of the kinetic theory of gases, which makes (b) a tautology. This means that if we read (ii) in the sense (a), disregarding what is in the brackets, we can see it as a *contingent a posteriori* thought, since it can be denied without contradiction, while if we read (ii) in the sense (b) it can be considered *necessary a priori*, since it cannot be denied without contradiction.

Consider now the sentence 'Water is H_2O.' I think Avrum Stroll was right when he noted that here the 'is' expresses *constitution*; the sentence more often means 'Water *is made* of H_2O' rather than 'Water is the same as (quantities of) H_2O.' (1996, 46 f.) However, this does not make a relevant difference for what I will try to say and contexts can lead us easily to read this 'is' as expressing identity.

As already noted (Appendix to Chapter II), the concept-word 'water' has two nuclei of meaning: a superficial one, that of an *aqueous liquid* (transparent, tasteless, odorless, etc.), and a deep one, a substance called by chemists *dihydrogen oxide* or H_2O (which includes much more than the simple chemical structure). This means that the complete sense of water must include the two nuclei. However, as in fact the presence of only one nucleus already allows us, in a proper context, to call the substance water, the most embracing criteria for the application of the general term 'water' demands sufficient satisfaction of the (summarized) inclusive disjunctive rule:

DR: (Water is an) aqueous liquid and/or (water is) dihydrogen oxide (H_2O).

Philosophers have created a pseudo-problem by insisting that the criterion of application of the conceptual word 'water' must be either aqueous liquid or dihydrogen oxide, as if it were a dilemma.[37]

Now, assuming that the 'is' is one of constitution and not of identity, the statement (i) 'Water is H_2O' in fact means: (ii) 'Aqueous liquid and/or dihydrogen oxide… is made of dihydrogen oxide.' Since it could be that water isn't made of dihydrogen oxide and only the first statement of the DR is true, it is possible for the whole statement to be false, which makes it a *contingent a posteriori* truth and not a *necessary a posteriori* truth, as Kripke would like it to be. However, as we will see in the next section, in some contexts statement (i) is rather seen as a *necessary a priori* truth.

28. Contexts of interest: no need for a necessary a posteriori

This double core sense of the general term 'water' helps to explain Saul Kripke's in my view as much insightful as illusory discovery of the *necessary a posteriori.* But in order to better understand the confusions involved, we need to add to the sentences the contexts in which they are spoken.

A first point to notice is that in the case of a sentences of the kind a = b uttered in different contexts we can *enhance* or *magnify* or *emphasize* its immediate (Fregean) perspectival sense that builds a derived thought (representing a sub-fact), or we can *enhance* or *magnify* or *emphasize* its

[37] For instance, A. J. Ayer in the first case and Hilary Putnam in the second. (See also Costa 2014, Ch. 3.)

mediated sense that builds the basal thought (representing a grounding fact).[38] Thus, in cases like 'Water is H_2O' we can emphasize the immediate core sense of the concept-word 'water' as an aqueous liquid or its mediated core sense as dihydrogen oxide. Here I need to speak again of the *contexts of interest* of the linguistic agents, meaning thereby contextualized practical aims from which we can infer what is meant.

Two contexts of interest are important regarding the main examples above: the *popular* and the *scientific* one. Thus, considering the sentence 'The morning star is the evening star,' we can contextually *emphasize* the derived thought composed by immediate senses (modes of presentation, identification rules) representing the external, phenomenally given objects, considering the difference between being the brightest star in the morning and the brightest star in the evening. If we do this, we leave the identity 'Venus = Venus' in the background. This can be the case, for instance, when contemplating the beauty of the starry sky at night and, after localizing the evening star, we tell a child that it is also the morning star. In this case, we think like Frege. We emphasize the different modes of presentation of the same object, a difference that as such represents nothing but an empirical sub-fact made by two different aspectual presentations of what we believe to be the same thing. We regard the thought that the morning star is the evening star as contingent a posteriori, since it mainly represents the sub-fact of the difference, although we are also aware that we are emphasizing the different ways by means of which the same thing presents itself to us.

Nonetheless, in a scientific context of interest, such as one in which astronomers use a telescope to study the surface of Venus, when they consider the sentence 'The morning star is also the evening star,' what they usually have in mind and emphasize is the numerical identity of the object of both modes of presentation. These are the mediated senses constituting the basal thought representing the grounding fact of the self-identity of Venus, which Kripke particularly emphasized in his writings. In this case, we read the statement as preferentially meaning the basal thought that 'Venus [in full] = Venus [in full],' which is a *necessary a priori* statement, since what we above all affirm is the tautological grounding fact that being Venus is the same as being Venus. It leaves the different guises of sense in the background, as secondary effects, insofar as we assume the truth of our scientific astronomical views.

[38] The concept of *emphasization* was fruitfully applied in Jürgen Habermas's excellent work on universal pragmatics (Habermas 1976).

Now, consider again the statement 'Water is H_2O'.[39] In a popular context of interest which arises when fishermen decide to dig a well to obtain fresh water for drinking and washing, this statement is read as emphasizing the sense of the word 'water' as a precious aqueous liquid (transparent, tasteless, odorless, drinkable… the popular nucleus of meaning), and it is for them a contingent matter that it is made of H_2O insofar as it satisfies their practical aims. Because of this, the statement is seen as *contingent a posteriori*, since it means 'This aqueous liquid is made of H_2O,' this expressing a derived thought representing a sub-fact that does not demand that water is necessarily H_2O, being deniable without contradiction.

On the other hand, when the context of interest is scientific, for instance, formed by chemists measuring the acidity of a sample of water, the word 'water' in the sentence 'Water is H_2O' can be read as emphasizing the sense of water as dihydrogen oxide (the scientific nucleus of meaning). In this case, the whole sentence is seen as preferentially expressing a thought representing a grounding fact expressed by the identity 'Water [H_2O] = H_2O [water],' which has the form $a = a$, that is, of a *necessary a priori* tautology based on our intuitive and scientific assumptions.

I think that philosophers like Kripke, by considering 'Water is H_2O' a *necessary a posteriori* statement, simply confuse (i) the *aposteriority* of the statement which emphasizes that water is an aqueous liquid made of H_2O with (ii) the *a priori necessity* of the statement that emphasizes the convention that water must be the same as H_2O, mixing the *aposteriority* of (i) with the conditioned *necessit*y of (ii).

A somewhat different emphasis can be found in the statement 'Heat is molecular movement,' here understood as 'Heat = molecular movement.' If we emphasize the ordinary immediate senses, the derived thought, the difference between heat2 (heat as it is normally felt) and the average kinetic energy of a gas, the emphasized sense or thought is *contingent a posteriori,* and the fact referred to is something learned by experience. This could be the case even using heat3 (temperature) as a fallible measure of average kinetic energy.

On the other hand, if we assume the truth of the kinetic theory of gases in a scientific context in which we are measuring temperatures, the statement can be understood as emphasizing the mediated sense of the identity expressible by: 'Temperature of a gas [-average kinetic energy] =

[39] The example was already considered in the *Addendum* of the Appendix to Chapter II in this book.

average kinetic energy [-temperature of a gas],' insofar as it is read as expressing the basal thought representing the grounding fact of an assumed identity, being therefore a (conditional) *necessary a priori* thought. In this reading, our conceptual rules for temperature and for average kinetic energy are blended into a single identification rule which assumes the kinetic theory of gases.

It seems to me that by considering identities of the kind *a* = *b,* Kripke misleadingly conjoined the *aposteriority* of the emphasized derived identity thought with the *necessity* of the emphasized basal identity thought, concluding that the identities between nominal and conceptual terms have a *necessary a posteriori* nature that is only metaphysically explicable. However, if these names or concept-words serve as rigid designators applying to the same entities in all possible worlds, this is explained by their assumed mediated senses, which are of the kind *a* = *a* (or a[b] = [a]b) and not only as *a* = *b* representing a difference. A Wittgensteinian therapist would conclude that in the considered cases Kripke was the victim of deep grammatical ambiguities. Finally, insofar as the terms *a* and *b* used in identity sentences are viewed as rigid designators unavoidably applying to the same ultimate object in all possible worlds where it exists, this is also only justified by the self-identity of a grounding fact.

29. Sense of a sentence: the thought

Now it is time to consider the sense of a sentence. Here is Frege at his best! He made the right decision in suggesting that the meaning of the whole sentence is the *thought* (*Gedanke*) it expresses. To reach this conclusion, he applied his compositionality principle: combined in the right way, the senses of the component terms constitute the sense of the whole sentence. If, for instance, in the sentence 'The morning star is a planet' we replace the description 'the morning star' with the description 'the evening star,' which is co-referential though having a different sense, the reference of the sentence does not change; but the sense of the sentence must change. Indeed, the sense of the sentence 'The evening star is a planet' is different. However, the only other thing that has changed is what we use to call *the thought* expressed by the resulting sentence. Consequently, the sense of a sentence must be the thought it expresses. (Frege 1892: 32)

The word 'thought' is ambiguous. One can use it to describe a *psychological process* of thinking, as in the utterance 'I was just thinking of you!' But it also seems to designate something independent of specific

mental occurrences – a content of thought – such as the thought expressed by the sentence '12 *x* 12 = 144' in the utterance: 'The sentence "12 *x* 12 = 144" expresses a true thought.' Frege had the latter sense in mind. In this usage, the word 'thought' means simply *what the sentence* (*statement*) *says*, which Frege conceived of as some sort of eternal (timeless) Platonic entity. A way to make the difference explicit would be to call the Fregean thought a *thought-content*. The terminology here counts because the word 'thought' is the only term in ordinary language that has a sense corresponding to more technical terms like 'proposition' or 'propositional content.'[40]

Frege has a criterion for deciding what belongs to a thought. For him, *everything that contributes to determining the truth-value of a sentence should belong to its thought*. Thus, using his own example, the sentences 'Alfred hasn't arrived' and 'Alfred hasn't arrived yet' express the same thought, since the word 'yet' means only an expectation regarding Alfred's arrival without contributing to the sentence's truth-value (Frege 1918: 64). The sentences 'The morning star is Venus' and 'The evening star is Venus' can be considered to express different thoughts because although the singular terms that make up these two identity sentences all refer to the same planet, they do this by means of different modes of presentation. That is, they make us follow different paths in determining their truth-value, or, as I prefer to think, they make us follow different associations of semantic-cognitive rules able to constitute correspondingly different verifiability procedures.

30. The thought as the truth-bearer

Another quite plausible Fregean thesis was that the *primary bearer* of truth is not the sentence, but rather the thought (proposition) expressed by it. I agree with this view. Although we can say that sentences, beliefs and even things and persons are true, they all seem to be true in a derivative sense.

Consider the cases of things and persons. A useful test to identify secondary uses is that when a word is derivatively used *we can replace it with a more appropriate word*. If we say that a diamond is false, what we mean is that it is only an *imitation* diamond: a *fake* or *counterfeit* of a real diamond that deceives us so much that we can think false thoughts about it. When we say that Socrates was 'true' as a person, what we mean is that

[40] As Tyler Burge wrote: 'the word "thought" is the best substitute for 'proposition' for the naturalness of its semantics within the scope appropriate to the linguistic philosophy.' (Burge, 2005: 227-8)

he was a *truthful*, *trustworthy* or *reliable* person, someone with *integrity*. But this is not always so. When we say that Sam's belief is true, although we secondarily mean that he has a subjective psychological attitude concerning his (dispositional) thought – of finding it true – we primarily mean that his thought is true in a Fregean sense.

One reason for preferring to say that the thought is the truth-bearer concerns the logical behavior of this concept. We deal with our concept of truth as an 'as if' directive idea, so that the real or actual truth-value of a thought is naturally conceived of as something *invariant*: if something is (really) true, it is always true; if something is (really) false, it is always false. Obviously, we can always err in *judging* and *claiming* something to be true (as *das Fürwahrhalten*) and can later discover it is false, and we can err in believing something to be false (*das Fürfalschhalten*) when it is actually true – this is often the case, and this possibility is inevitable, due to our inherent epistemic fallibility. But when we discover our error, we correct ourselves, in the first case not by claiming that the thought was previously true and now has become false, but by saying that it was *always false*, and in the second case we correct ourselves not by claiming that the thought was previously false and now has become true, but by saying that it was *always true.* What changed was our truth-claim expressing our judgment, not the truth-value. Moreover, it is fundamental to perceive that our inherent fallibility in holding thoughts to be true does not affect the invariability or immutability of the truth-value of the thought or proposition in itself. It must be so because it is beyond our fallible capacities to know with absolute certainty whether we have achieved this ideal, if we have indeed achieved it. This is how the logical grammar of our concept of truth works (and, beyond this, the grammar of our own concept of knowledge). If one wants to change something so fundamental, then to prevent confusion one should invent new terms instead, like 'hturt' and 'eslaf'.

Now, if the actual truth-value is immutable, its truth-bearer must also be unchanging, able to remain the same in order to retain this same truth-value independently of the time or place where we discovered it. Indeed, for Frege a really true thought remains true forever, just as a really false thought remains false forever. These entities are even abbreviated as 'truths' and 'falsities' respectively. Thus, it is deeply ingrained in our conceptual grammar that the entity that can be primarily called true or false must remain the same and possess the same truth-value so that what may change is only our cognitive grasp of it, our believing in its truth-value (*das Fürwahrhalten*). If this is so, then only the thought has the necessary *stability* to be the archetypical truth-bearer; for a thought is,

according to Frege, unchangeable and eternal (a-temporal), being eternally (a-temporally) true or false independently of our grasping (*fassen*) it.

Consider now the case of sentences as candidates for truth-bearers. Ambiguous sentences can express different Fregean thoughts, such as 'John saw the man on the mountain with a telescope.' In this case, the truth-value of the thought will be able to change according to the different thoughts or interpretations that we assign to the sentence. But if the truth-bearer were the sentence, the truth-value should remain the same, which cannot be correct. This is obvious in the case of indexical utterances like 'I am in pain,' which has different truth-values depending on the speaker.[41] The same sentence can change its sense-thought when uttered by different persons, and even when uttered by the same person at different times; correspondingly, what may change with the change in thought is the truth-value. Hence, thoughts and their truth-values are *co-variant,* while sentences and their truth-values are not, which leads us to the conclusion that the primary bearer of truth-value must be the thought or proposition.

One could suppose that perhaps the *sentence-token* would be the truth-bearer, since it would be a different one depending on the time and place of the utterance, changing with the truth-value. However, we still have cases in which different sentences (token or not) say the same thing – express the same thought – in this way preserving the same truth-value. Consider, for example, the following statements, 'It is raining,' 'Il pleut,' 'Es regnet,' 'Chove'… uttered in the same context. They all say the same thing, express the same thought, and all have the same truth-value, while their sentence-tokens are quite different. Indeed, the only justification for insisting on the immutability of the truth-value of these four different sentence-tokens (and types) is that their primary truth-*bearer* is the thought expressed by them, since what they say – their senses, their thoughts – is what remains the same. Finally, this is the case not only for indexical sentences but also for eternal sentences with the same content, though expressed in different languages.

Likewise, beliefs, understood in a psychological sense, can only be derivative truth-bearers: if someone who believes something dies, his psychological belief also ceases to exist. Consequently, the truth-bearer must be the content of his belief. It must be his belief-content and not his belief in a dispositional psychological sense, since only the belief-content

[41] For Frege, in the case of indexical sentences, the context of the utterance belongs to the expression of thought. See also addendum of the Appendix to Chapter II, sec. 8.

isn't transitory. But this is so only because we understand the belief-content as the same as a Fregean thought, a propositional content.

The core of the foregoing arguments can be summarized as follows: thoughts and their truth-values are not just invariantly related; when thoughts vary, they maintain *a relationship of co-variance with their truth-values*. This relationship is missing in the relationships between sentences or psychological beliefs and their truth-values. Because of this, the proper bearer of truth must be the thought (proposition, propositional content, belief-content), not the sentence or some personal psychological disposition to agree on a truth-value.

31. Facts as true thoughts?

As already noted, Frege also proposed that what we call a *fact* is the same thing as a *true thought*, because when a scientist discovers a true thought, he claims to have discovered a fact. As he wrote:

> 'Facts! Facts! Facts!' exclaims the researcher of nature, when he wants to proclaim the need for a secure basis of science. What is a fact? A fact is a thought that is true. (1918: 74)

Indeed, when we say 'John stated several relevant facts in his speech,' we are speaking about facts as true thoughts. However, there is no warrant that this is not a derivative use of the word 'fact.' A researcher of nature can well exclaim 'Facts! Facts! Facts!' understanding by a fact simply what *corresponds* to the true thought, namely, some objectively given tropical arrangement. After all, it seems natural to think that if someone discovers a true thought, it is because he has *a fortiori* discovered the fact corresponding to it.

A more decisive argument against thoughts as true facts came from J. L. Austin, who made it clear that Frege's identification does not resist all linguistic replacements (1990: 170-171). If the sentence 'What he affirms is true' had the same sense as 'What he affirms is a fact,' then the replacement of 'what he affirms' with 'his affirmation' should be allowed without any change of sense. But, 'His affirmation is true' preserves the meaning, while 'His affirmation is a fact' makes sense only as a meta-linguistic sentence referring to the occurrence of his affirmation, and not to the content of the affirmation itself. The reason for this can only be that the true content of an affirmation – the Fregean thought – cannot be properly identified with a fact.

The main reason why Frege believed that a fact is a true thought is that he advocated a conception of truth as *redundancy*, rejecting the

correspondence theory. However, on the one hand, his arguments against correspondence theory (1918: 59-60) are unconvincing.[42] On the other hand, correspondence theory remains the *prima facie* most plausible view. It is the most natural and historically influential conception of truth, suggesting that propositions or thoughts are true when they correspond to facts as arrangements of elements in the world (Rasmussen 2014; Vision 2004). Moreover, the view of truth as *correspondence* is commonsensical, agreeing with our methodological principle of the primacy of common knowledge. Because of this, I will defend this theory in the last chapter of this book.

Finally, I think I have found a plausible way to explain why some are tempted to say that facts are true thoughts. It seems that the source of confusion resides in a persistent ambiguity of our own natural language. Dictionaries in very different languages present us a variety of trivial meanings for the word 'truth.' However, two general meanings are almost invariably emphasized. I call them: *thought-truth* and *fact-truth*. Here are their definitions, according to the best dictionaries:

(a) *Thought-truth*: Truth as consisting of things being as we believe they are, as conformity or accordance or correspondence of the thought with the fact it represents.
(b) *Fact-truth*: Truth as the actual, real, existing fact in the world.[43]

It is regarding the philosophically most proper sense (a) that we have singled out the thought as the primary bearer of truth. This usage is shown clearly in sentences like 'His words are true,' 'Tell me the truth.' In the factual sense (b), we single out facts in the world as secondary truth-bearers in the sense of *being real*, and we use sentences like 'The mentioned occurrence was true (was real),' 'We are searching for the true facts (the real facts),' 'The truth (the fact) is out there.' The possibility of

[42] According to his main argument, if you say that the truth of *p* is its correspondence with reality, you need to admit that *p* must have the property φ in order to be true by correspondence with reality, and that to have the property φ in order to be true by corresponding with reality will demand the property φ' and so successively. The answer (already given by Aquinas) is that to say that *p* is true by corresponding to reality, and to say that *p* has the property φ due to being true by corresponding to reality are *one and the same thing*; consequently, N is redundant. (*Cf.* Künne 2003: 129-133).

[43] For instance: 'truth (principle): that which is true in accordance with the fact or reality'; 'truth (fact): the actual fact about the matter'… (*Oxford-Cambridge Dictionary*).

more adequate semantic replacements indicates the derivative character of fact-truths.

As we have already seen, there are good reasons to think that sense (a) is primary while sense (b) is derivative, since in this last case we can replace the word 'truth' with more adequate ones like 'reality,' 'existence,' 'actuality'… Anyway, 'truth' is very often used not only as 'correspondence with facts' but also replacing 'an existing fact in the world.' Thus, we can easily be misled by some extraneous motivation and confuse the two usages, mistakenly concluding that facts are true thoughts. This is what seems to have originated Frege's confusion, giving us another example of *equivocity* as a common way of transgressing the internal limits of language (Ch. III, sec. 11).

32. The thought as a verifiability rule

As the application of the ascription rule (sense of the predicate) is subsidiary to the application of the identification rule (sense of the nominative term), the rule for applying the singular sentence (its sense or thought) can be seen as an association of semantic-cognitive rules. Ernst Tugendhat has identified this association with the *verifiability rule* in the case of the singular predicative statement (1976: 259, 484, 487-8), which implies the suggestion that this view can be generalized to all meaningful statements (See 1983: 235-6). Indeed, if the thought is an association of rules, then what results from such an association – the verifiability rule – must also have the character of a rule, even if it isn't something previously conventionalized. Combining this with our acceptance of the correspondence view of truth and our salvaging of the fact as the universal truth-maker, this means that the thought should be a kind of associated or combined semantic-cognitive rule – a verifiability rule – whose function is to make us aware of a corresponding fact to which it is applied.[44]

This reasoning unavoidably leads us back to the controversial idea of 'verificationism,' more precisely (and still worse) to *semantic verificationism*: the doctrine first proposed by Wittgenstein, according to which *the (cognitive, informative) sense of a sentence is the rule or method or procedure used in its verification* (1980: 29). As it is well-known, Wittgenstein's idea was soon appropriated by the philosophers of logical positivism. However, after varied attempts to give it a precise formulation, it was in the end abandoned due to strong criticism, internal and external to the logical-positivist circle, which led to it being considered by many as

[44] See Tugendhat's verificationist correspondentialism in 1983: 235-6.

unsustainable. This is presently the received view, even if sophisticated philosophers have never really abandoned the idea that some form or other of verificationism is indispensable (*Cf.* Misak 1995). Indeed, in the next chapter of this book I intend to offer replies to the main objections that philosophers have made against semantic verificationism, showing that these objections were not directed against its correct form, but rather against a straw-man called the 'principle of verifiability,' as it was wrongly construed by logical positivists.

I am introducing semantic verificationism in this chapter speculatively, as an alternative and in fact as the most natural way to analyze Frege's discovery of the thought as the cognitive sense (epistemic value, informative content) of a sentence. Now, suppose that the combined semantic-cognitive rule that constitutes the thought as expressed in an assertoric sentence is its verifiability rule, as complex as it may be. Then the verifiability rule in itself is the most proper truth-bearer. Then, if we show that this verifiability sense-thought rule is effectively applicable to the expected fact, this makes the rule true, which allows us to say derivatively that the sentence expressing it is also true. If, on the other hand, we show that this thought-sense-rule, though conceivable, isn't effectively applicable to the expected fact, this makes it false and likewise the sentence expressing it. Moreover, if we cannot formulate a verifiability rule able to be at least in principle applicable to the fact, if we cannot even conceive its application, we must conclude that the declarative sentence is devoid of meaning, devoid of sense or thought, even if it may in some cases seem to have meaning.

I think that this way to understand the truth of a thought is in line with Frege's remark that although he regarded truth as the property of a thought, it does not seem to be a property in the usual sense of the word (Frege 1918: 61). Indeed, truth does not add anything to the combined cognitive rule called 'the thought,' except something dispositional, namely, its effective applicability as a verifiability rule in the appropriate context for its application. Moreover, the proposed identity between the Fregean concept of thought and the concept of a verifiability rule is also supported by the Fregean proposal that the identification criterion for what belongs to a thought is that it must have at least some role in the establishment of the thought's truth-value.[45]

[45] Nonetheless, there is an at least seemingly alternative way to understand the property of effective applicability of the verifiability rule, which is to identify it with the *existence* of the fact. To reach this conclusion, we need only consider that the existence of an object (an independent cluster of compresent tropes) is the higher-order property of effective applicability of an identification rule expressed

33. Frege's Platonism

It is important to remember that for Frege thoughts and the senses that compose them are Platonic entities belonging to a third ontological realm, which is neither psychological nor physical (Frege 1918). For him, taking (a) the criterion of *objectivity* as being inter-subjectivity and independence of will, and taking (b) the criterion of *reality* as existence in space and time, we combine them in order to get three ontological realms:

1. Realm of the *objective and real*
2. Realm of the *subjective and real*
3. Realm of the *objective but non-real*

The first realm is that of physical entities such as concrete objects, which are *objective and real*. These entities satisfy criteria (a) and (b): they are objective, since they are interpersonally accessible and independent of our will, and they are real since they are located in space and time. The second realm is that of psychological entities, mental states that he calls representations (Frege uses the word '*Vorstellungen*' in a way that could be easily translated as *qualia*). These entities satisfy criterion (b) but not (a): they are *subjective and real*. By not being interpersonally accessible, they are subjective and often dependent on the will. However, they are still real, because they are in the mind and, consequently, in time and (we can add) space. There is, finally, a third realm, that of thoughts (propositions) and their constitutive senses. This realm satisfies criterion (a) but not (b). For Frege thoughts are *objective but not real*. Thoughts are objective, because, true or false, they are always interpersonally accessible: we can all agree, for example, that the Pythagorean Theorem expresses a true thought in Euclidean geometry. However, this third realm of thoughts is not real, because according to him thoughts are abstract things that cannot

---

by a nominal term, and that the existence of a property – a dependent property-trope – is the higher-order property of effective applicability of the ascription rule of a predicative expression. If we accept this, then by symmetry the existence of a singular fact should be the higher-order property of effective applicability of the verifiability rule of the singular declarative sentence to which it applies. It seems that we could say, in an almost Hegelian fashion, that existence is the truth of the concept, while the truth is the existence of the thought… We have here two alternative understandings of the property of effective applicability of a verifiability rule, what generates a dilemma that will only be solved in the beginning of chapter VI.

be found in space or time. Thus, the thought (the sense) of Pythagoras' theorem is objective but non-real.

There are, however, problems. One of them, noted by Frege, is that although for him thoughts are eternal (timeless), immutable, forever true or false, and not created but only *grasped* (*gefasst*) by us, they must have some kind of *causal effect*: they must be able to cause our grasping them in order to make judgments and act in the external world (Frege 1918: 77). How this interaction with something non-spatiotemporal is possible remains an unexplained mystery.

Frege was aware of the difficulties, but the main reason why he felt he had to introduce this third realm of thoughts is that thoughts are interpersonally accessible, that is, they are objective, which makes them able to be communicable. Representations (*Vorstellungen*), on the other hand, are rather subjective psychological states that can vary depending on personal psychology and according to him could never become interpersonally accessible and therefore are not communicable. Thus, for him the right way to explain how it is possible that we are able to share the same thoughts in conversation is to strictly distinguish thoughts from mere psychological representations, placing them in a supposedly shareable Platonic realm. In addition, if thoughts were on the level of representations, they would be dependent on changeable personal psychology and would lack their required stability as truth-bearers.

34. Avoiding Frege's Platonism

Despite the above-suggested arguments, few today would accept Frege's appeal to Platonism. After all, the Fregean form of Platonism not only commits us to an infinite multiplication of objective entities (all the infinite variety of true and false thoughts and their constitutive senses) but also seems to lack intelligibility. The price that Frege was willing to pay in order to avoid psychological subjectivism seems too high for us today.

In my judgment, if we understand senses as rules, which usually are implicitly established conventions or something derived from them, there is a clear way to bring the empiricist view of thoughts as having a psychological-empirical nature in line with the view that as truth-bearers they must have stability and the possibility of being communicated. In order to establish this conclusion, I want to apply again the same strategy inspired by the ontological particularism of English empiricists, which I used in the construction of universals by means of tropes.[46] This is

[46] See Appendix to Chapter III, sec. 2.

understandable since according to trope ontology, a thought should be made up of, at least dispositional, internal tropes: the mental tropes constitutive of some conventionally grounded verifiability rule whose application is at least conceivable. In order to accomplish this, I need only show that something like Fregean Platonic thoughts (objective non-real truth-bearers…), which I call *f-thoughts* ('f' from Fregean) can be defined in terms of psychological (real and subjective) *p-thoughts* ('p' from psychological), though typically based on intersubjective linguistic conventions. In other words, I suggest that we can warrant the existence and stability of f-thoughts without hypostasizing them as Platonic entities and even without resorting to classes of p-thoughts if we replace them with what I call *extensible thoughts* or *e-thoughts*. We can do this by means of the following disjunctive definition, which is as simple as it is efficacious:

> An e-thought (*Df*) = a given tropical p-thought *X** (used as the model) embodied in some mind or any other tropical p-thought *Y* qualitatively identical to *X**, embodied in the same mind or in any other mind.

The e-thought is our empiricist version of what Frege should have meant with his f-thought (objective non-real thought). The p-thought *X** can be any *X* thought that someone decides to use as a model. The aim of this definition of an e-thought is that any supposed f-thought is reduced to mental p-thoughts without depriving it of its epistemic objectivity (mainly inter-subjectivity) grounded on conventional rules, along with its expected stability or immutability. This procedure works at least insofar as my criticism of the private language argument is acceptable, though I have no doubts about this (See Ch. III, sec. 13).

The so defined e-thought – which is the same as a verifiability rule, a tropical *thought-content* or simply a *proposition* – though usually distributed across space and time, doesn't need to have any particular spatiotemporal location and can be seen as the most proper truth-bearer. For example: the e-thought or e-thought-content or e-thought-content-rule expressed by the sentence 'The Eiffel Tour is made of iron' can be instantiated as the p-thought that I have in mind when writing this sentence. However, it can also be instantiated by, say, the p-thought that you have in mind when you read it, such as by any qualitatively identical p-thought that I, we, or any other person can have at any place or time, insofar as it is considered an f-thought, namely, a model for any qualitatively identical p-thoughts. Characterized by the disjunction between qualitatively identical thoughts embodied in individual minds, the e-thought is apt to be regarded in abstraction from any *particular* human

mind that causally instantiates it. This is what really occurs when we think an f-thought, and it is this abstraction from singular human minds resulting from the spreading character of the real thought-contents that gave Frege the impression that he had found a Platonic entity outside of space and time.

As with model-tropes in the construction of universals, it is not necessary to have only one particular model as the object of interpersonal consideration. To the contrary, what we need to do is simply to single out the first thought given to us by memory and use it arbitrarily as a model: first the one, and then any other that we recognize as being precisely (qualitatively) identical to the first, and we can choose any of them as a new model. In some way language is only the vehicle of communication that allows the reproduction of qualitatively identical psychological p-thoughts in the minds of hearers, insofar as they are rooted in the usually implicit interpersonal conventions we have attached to their semantic components. Since p-thoughts are tropes, we have simply applied to p-thoughts the same strategy we applied to singular tropes, as we needed to construct universals based on them. The e-thought verifiability rules are p-thoughts read as universals.

With the help of the above definition, we avoid not only appealing to psychologically specific occurrences of thoughts, but also the most expected alternative, which would be to explain one e-thought in terms of a sum or set of p-thoughts qualitatively identical to each other. This could lead us not only to the problem of defining sets, but also to the problem that sets and sums have or could have size, while thoughts cannot. If an e-thought were a set of p-thoughts, even if considered an open set, it would at the ontological level grow ever larger, the greater the number of people there were who grasped it.

Under the proposed definition, in order to exist, an e-thought must always have at least one psychological occurrence. The e-thought is not less psychological than any p-thought, since it cannot be considered independently of its instantiation *in at least one mind*. This means that when we say that we both had the *same* idea, or the *same* thought, this is merely a manner of speaking. What we really mean is only that there is a qualitative identity between the (tropical) psychological verifiability p-thought-contents rules that we have respectively instantiated in our minds. We share the e-thought in the sense that we instantiate qualitatively identical p-thoughts. This has the advantage of bringing Fregean thoughts out of the ethereal Platonic heaven back to the concrete psychological realm without making any serious commitment to the transient psychology of individual minds.

This understanding of the true nature of thought-contents explains something that Frege was unable to explain satisfactorily, namely, why and how they may have causal powers. Since as an open disjunction of p-thoughts, e-thoughts only exist as psychological instantiations of p-thoughts, this enables them to play a causal role: they can cause other psychological states and, combined with desires, human actions and their effects in the external world.

At this point one could raise an objection of *multiple realizability*: the same p-thought could be differently realized in different human brains, making the qualitative identity of p-thoughts impossible. I agree with the very probable multiple realizability of p-thoughts but disagree that this makes their qualitative identity impossible. There is no reason why we cannot present things that can be considered qualitatively identical on a linguistic or even psychological level and different on a neurophysiological level, in the same way as different devices can have different internal mechanisms and perform exactly the same tasks.[47] Moreover, my suggestion is that e-thoughts are constituted of p-thoughts that are internal tropical verifiability rules, which although complex, ramified and variable, are also able to be satisfied by foreseeable independent tropical configurations.

In my judgment, one of the most unyielding and deceitful philosophical errors in ontology has always been *seeing numerical identity where there is only qualitative identity*. It is true that we can ask for *the* meaning of the general term 'chair' using the definite article 'the' in the phrase '*the* chair.' But this is only a linguistic device that changes nothing! In a similar way, we can speak of *the* geometrical form of circularity, and of *the* number 2 in the singular… But this is just for the sake of simplicity of expression. What we are ultimately able to have in mind in all these cases are occurrences of qualitatively identical meanings, that is, of qualitatively identical concepts of chairs, circles, and cognitive arithmetical concepts of duality, and not something more, since we don't need something more to get something more.[48] In the same way, we can talk about the thought

[47] As T. W. Polger has shown, in order to illustrate the flaw of the multiple realizability argument, we can explain how a carburetor has the function of mixing fuel and air for a combustion engine; but it is a multiply realizable device: it can be made of various different materials with various designs, provided it functions properly. (2004: 19-20).

[48] The phrase is from Murray Gell-Mann. Against this, however, one could ask: haven't we learned that geometry deals with perfect circles and that arithmetic deals with entirely abstract numbers? Take the case of circles. The answer is, of course, in the negative, because we can make a new circle more perfect than the last one, and another even more perfect, and this process can continue indefinitely.

expressed by '12 *x* 12 = 144,' but if we do not intend a specific occurrence of this thought, we are only referring to *some* occurrence, but without taking into account or having to specify *which* occurrence and *in what* mind. We speak in the singular of the thought that 12 *x* 12 = 144 for reasons of simplicity.

The adoption of the definition of e-thoughts proposed above, which is easily generalizable to all kinds of Fregean senses, seems to me the only plausible abstraction we can arrive at without committing any of various forms of reification that have infested ontology throughout its long history.

At this point, a stubborn Fregean defender can still ask: how is it possible that the psychologically dependent definition of e-thoughts suggested above could be able to ensure the objectivity of e-thoughts, their interpersonal accessibility or communicability? As we saw, Frege concluded that if we regard thoughts as psychological representations, as is the case with p-thoughts, they would unavoidably be subjective, and we could not compare them with each other. However, it still seems clear that Frege was too hasty when he admitted that his f-thoughts belong to a third realm of Platonic entities. One could note that there is no doubt that what Frege calls representations (*phenomenal* mental contents) have in fact possibilities of interpersonal communication, even if limited.[49] But much more important is something that Frege hasn't considered at all, namely, that senses and e-thoughts, without being Platonic entities, could be understood as rule-complexes built upon adequate associations of interpersonally accepted conventions established with the help of public signs that are communicable precisely because of their grounding interpersonal character. That is, because e-thoughts are verifiability rules rooted in linguistically shareable interpersonal conventions, they can well be able to satisfy

---

The perfect circle is like the actual infinite: it does not exist. It is nothing more than a projection of our awareness of the possibility of making increasingly perfect empirical circles without any conceivable end. Geometry does not work with actual perfect circles, but with potentially perfect circles.

[49] Against Frege, we could hold that to some extent even imagetic representations can be expressed through language and by its means could be subjectively identified and re-identified as being the same (e.g., a police sketch or a Photofit). It is true that a mental state that only one person is capable of having, for instance, a sort of epileptic aura, is not communicable, except indirectly, metaphorically. But it seems very plausible that typical mental states, such as feelings, images, sensations, are things that all of us are able to communicate and learn to identify in ourselves through induction by exclusion, added to induction by analogy and reinforced by a great variety of interpersonally accessible physical states strongly intermingled with them (*Cf.* Ch. III, sec. 8; See also Costa 2011, Ch. 3).

Frege's demand for objectivity as interpersonal accessibility followed by the possibility of communication and truth-evaluation.

It may, at first sight, seem implausible that language is capable of repeatedly being reproduced in other minds and even in the same mind with the same subjective pattern, the same thought-content, the same recognizable instantiation of an adequate association of conventionally established semantic-cognitive rules attached to our words. However, compare by analogy this case with that of genetic information able to endlessly reproduce the same characteristics in successive biological individuals.[50] Why cannot the conventions and ways they can be combined in the constitution of p-thoughts do a similar job, even if only inferentially? More than this (and probably also in the case of genetic information), it is easy to suppose that there are corrective mechanisms able to interpersonally and intra-personally impose a limit on divergence from conventionalized standards (See Ch. V, sec. 11). There is no reason, except an anti-empiricist bias, to think that things could not be that way.

Finally, let us apply to e-thoughts John Searle's important distinction between what is *ontologically objective/subjective* and what is *epistemologically objective/subjective* (Searle 1999: 43-45). Searle noted that we have a strong tendency to take what is epistemologically subjective for what is only ontologically subjective. However, something can be ontologically objective – for instance, 'How justifiable was the First World War?' – without ceasing to be epistemologically subjective, because it is not easy to reach a common agreement about this issue. In contrast, a phenomenon can be ontologically subjective without ceasing to be epistemologically objective – for instance, the stabbing pain caused by a seizure of acute pancreatitis – because everyone (doctors and patients alike) will agree on the form and existence of this pain, even if the patient alone knows exactly how it feels.

Something of the kind can also be said not only of Fregean subjective mental representations, but also of e-thoughts. In themselves they are *ontologically subjective*, since we admit that they are psychological events instantiated in one mind or another. But even so, they do not cease to be *epistemologically objective*, since we are capable of interpersonally agreeing about them and their truth-values. We can agree that an objectively

[50] Biological mutations are accidents whose occurrence should be evolutionarily *calibrated*. Species are only likely to survive if they can mutate to the right degree in the right period of time in order to adapt to environmental changes. Too many mutations, as well as too few, would be dangerous for species survival. It seems possible that an unchanging species with no mutation is conceivable, but it would be unable to adapt to changing external conditions.

assertoric sentence like 'The Eiffel Tower is made of metal' expresses a true e-thought that is epistemologically objective, even though as an e-thought ontologically subjective, since it is distributed among the minds of those who think it. Like any e-thought, it remains epistemologically objective, given that it is grounded on conventions associating words with things in the world, which makes it fully measurable and communicable. An arithmetical sentence like '2 + 3 = 5' is epistemologically objective (since we are all able to inter-subjectively agree on its truth-value), but it also expresses an ontologically subjective e-thought, and as I tried to show in speaking of numbers, it seems to be a thin kind of tropical arrangement sustained by lower-order tropes. On the other hand, a sentence like 'Love is the Amen of the universe' (Novalis), unlike an e-thought, has no truth-value. It is only suggestive and expressive. Like poetry, it is based on non-conventional subjective coloration, being susceptible only to emotive-aesthetic appreciation with differing degrees of subjective interpersonal agreement.

Regarding ontology, Frege was no exception. Like Husserl, Bolzano and several other continental philosophers of his time with mathematical training, he believed that the ontologically subjective character of psychologically conceived thought-contents would inevitably be condemned to epistemological subjectivity. But this was a mistake.

35. Further ontological consequences

Our ultimately psychological reformulation of Fregean thoughts has some interesting ontological consequences. If the thought of the Pythagorean Theorem isn't an eternal (timeless) entity belonging to a Platonic realm, always true or false, where and when does it exist? The answer is that if there is at least one occurrence of its thought or any other qualitatively identical occurrence, regardless of the bearer, something like the Pythagorean theorem acquires an existence dependent on minds. It is not an existence dependent on any of the many particular minds that will eventually think it since it would continue to exist without having been thought by this or that particular mind. In fact, since this thought has been thought by both you and me and certainly by many others in the past, its existence must be *spread over space and time*. It must be distributed over the space and time occupied by the heads of mathematicians starting with Pythagoras himself and perhaps ending in the head of some cognitive being at some unknown future time. This is what gives the impression that *the thought* is something abstract, beyond the psychological realm.

Another consequence of the proposed view is that unlike the Platonic entity that Frege called a 'thought,' our e-thought of the Pythagorean theorem did not in fact exist before Pythagoras thought it for the first time (supposing he was the first), and will cease to exist if it ceases to be thought by anyone. The Pythagorean theorem certainly exists, has existed and will continue to exist in the sense that it is thought, has been thought and will probably be thought in the future, referring to occurrences of this thought, but without having to take into account who thinks it.

One could object that this result sounds strange: it seems that the Pythagorean Theorem applies independently of minds. However, this strangeness can be softened by the fact that nobody can truly deny it. One cannot have the true thought, 'The theorem according to which the sum of the squares of the shorter sides of a right triangle equals the square of the hypotenuse has been thought in the past and now is no longer thinkable.' And the reason is that this judgment will already be an occurrence of the thought of the Pythagorean Theorem and insofar will falsify what it states. Anyway, the conclusion remains that the e-thought of this theorem would not have come into existence if nobody had ever thought it. Putting this more incisively: it would not exist in a world without cognitive beings.

The last remark suggests the following objection. Imagine a possible world *Ww* similar to ours, with planets, stars, and galaxies, but without any cognitive being. In *Ww* the e-thoughts that there are planets, stars and galaxies could not have been thought and, e-thoughts, being the primary bearers of truth, could not be true. Nevertheless, it seems very reasonable to think that in this world the fact that there are planets, stars and galaxies would *still be true, even though there would be no cognitive Beings to think this.*

It seems to me that the right answer to the strangeness is that here we are again victims of a confusion between thought-truth and fact-truth. As we saw, the first is the truth applied to the primary bearer of the truth, which is the e-thought, while the second is a derivative but very common application of truth to the real existing thing or fact in the world, as a secondary bearer of truth, meaning a real thing or fact. Indeed, that there would be planets, stars and galaxies in a mindless world would still be *true as a fact* in *Ww*. Hence, the applicability of the Pythagorean Theorem would still be a *fact-truth* in *Ww*, even though neither their e-thoughts nor their truth in the form of correspondence would exist. The flexibility of natural language has once again misled us.

Still another objection that could be made against the idea that the bearers of truth are non-Platonic e-thoughts is the following. Many truths have been discovered. Pythagoras is credited with discovering the theorem

that bears his name; Archimedes was one of the discoverers of the law of the lever, according to which magnitudes are in equilibrium at distances inversely proportional to their weights. However, if something is discovered, then logically it must have existed before being discovered. Consequently, the above-described thoughts must already have existed before their discovery.

Again, the answer is that this naïve objection results from a confusion between the thought as the primary bearer of truth on the one hand, and the fact as a derived bearer of truth on the other. This is clear in the case of typical empirical truths. That the law of the lever was always applicable in principle is surely true. However, this is only a general *fact-truth*! Its thought-truth was only part of the empirical (mental) world after scientists like Archimedes conceived it. Similarly, common sense tells us that the fact expressed by the Pythagorean Theorem must always have existed. However, our e-thought of it only came into existence after the theorem was thought by Pythagoras and since then has been thought by many others. Real facts, on their turn, as long lasting as they may be, are not the primary bearers of truth, but rather their truth-makers or verifiers. They exist independently and are said to be true only in the derived sense (b) of fact-truths, not in the sense (a) of thought-truths. They are what occurrences of their thoughts represent. Hence, in the most proper and demanding sense, no truths or falsehoods would exist in a world where there were no minds to think them. The most we could think of in this direction is to say that if the law of the lever were thought in *Ww*, it would be recognized as true.

An e-thought that has never been thought does not exist and thus cannot be true. The same holds for falsehoods. Consider the thought 'The Colossus of Rhodes is floating in the Sargasso Sea.' In all probability this thought has never been thought before the present moment. But the moment we think that it has never been thought before, we are already thinking it, and we can even attribute falsehood to it. Even the e-thought 'The world could exist, even if there were no minds to think about it' is only a true thought insofar as there are minds to think it.

36. A short digression on contingent futures

Before we finish, it is interesting to examine the Aristotelian problem of contingent futures in the light of our conclusions (1984, vol. 1, *De Interpretatione*, sec. 9). According to a plausible interpretation of Aristotle, the following argument is valid:

Argument A
1. Necessarily, it is true or false that there will be a sea-battle tomorrow.
2. If (1) is true, then the future is predetermined and there are no chance events.
3. Therefore, the future is fixed and there are no chance events.

It seems that for Aristotle this conclusion would be unacceptable, because if the future were predetermined, then there would be no chance events, and if there were no chance events, there would be no free will. Hence, according to a traditional interpretation, he thought that although this argument is sound, premise (1) is false because it exemplifies the principle of bivalence, and the principle of bivalence – according to which any significant proposition is either true or false – is not applicable to future events (only to present and past ones).

I cannot agree with this conclusion, since I believe that we should preserve a strongly understood principle of bivalence for e-thoughts.[51] But premise (1) can be questioned from a different perspective. Suppose, first, that *outside any context* we consider the e-thought expressed by the sentence 'There will be a sea battle tomorrow,' which we can abbreviate as $\vdash p$. Is this statement true or false? The answer is the following: if taken literally, $\vdash p$ is unable to express any e-thought because a verifiability e-thought rule is something to which we must possibly attribute a truth-value. Normally 'There will be a sea battle tomorrow' is an incomplete indexical statement, so that without any further contextual information we are totally at a loss for the task of associating *p* with any appropriate truth-maker in order to assign it a truth-value.

Moreover, one could argue that the sentence $\vdash p$ (as much as $\vdash \sim p$) is misleading and causes confusion, like argument A, because $\vdash p$ only *seems* to express cognitive thought-content. The reason for this is that $\vdash p$ is very easily confused with the meaningful sentence $\vdash p^*$: '[It is *likely* that] a sea-battle will take place tomorrow,' stated when *there are reasons* to think so. For example: having broken Japanese naval codes and having lured the Japanese fleet into an ambush at Midway, the Americans already knew on the night of June 3, 1942, that on June 4 there would almost certainly be a major naval battle. The sentence $\vdash p^*$ is easily confused with $\vdash p$, because $\vdash p^*$ *is* almost always *abbreviated* as $\vdash p$: 'A sea-battle will take place tomorrow.'

[51] I mean a principle of bivalence understood as a different formulation of the principle of non-contradiction.

For example: suppose that American Admiral Nimitz had said on June 3:

Tomorrow there will be a sea-battle.

Everyone would understand that he was saying that all the factual evidence was leading to the conclusion that the expected battle would begin on June 4. This probability – made explicit or not – is in this case objectively measurable in terms of verification by actual empirical evidence, so that the assertion $\vdash p^*$ expresses an e-thought that is held to be true, for it is true that, with the information already available, it was very probable that a sea-battle would occur the next day. Indeed, the utterance 'It is likely that a naval battle will take place tomorrow' could be regarded as *definitely true* on the night of June 3, 1942, without violating any principle of bivalence!

Suppose now, by contrast, that I am standing on the calm beach of Praia Bonita in Northeastern Brazil, looking out across the Atlantic Ocean, and without any reason I say $\vdash q^*$: 'A naval battle will take place in this region tomorrow,' meaning by it 'It is likely that a naval battle will take place in this region tomorrow.' This statement can be regarded as *definitely false*, since there are many different reasons to believe that this kind of event is extremely improbable in this region and at this time.

The conclusion is that in the absence of a context (and not in the above senses of $\vdash p^*$ or $\vdash q^*$), the statement $\vdash p$ would be a *linguistic bluff* devoid of any meaning or justification. Aristotle would be right in rejecting the application of the principle of bivalence to it, not because this principle has exceptions, but simply because it expresses no e-thought, no proposition, no verifiability rule. All that this sentence does is to induce us to imagine a naval battle that takes place tomorrow, as if there were hidden verifiability criteria. However, insofar as no context is furnished, no real criteria can be given. Statements like $\vdash p^*$, $\vdash \sim p^*$ and $\vdash q^*$, on the other hand, aim to say something probabilistic about the future that can be confirmed and made true by criterial reasons already found in the present. But from such statements premise (2) and the conclusion (3) of the argument A do not follow, because all that such statements can warrant, if true, is the inductive probability of a sea-battle.

The upshot is that the metaphysical riddle about contingent futures can be eliminated if we consider with enough care what we are really able to mean by affirming e-thoughts regarding the future.

37. Conclusion

My first aim in this chapter was to insert in the framework of Fregean semantics the results of my reconstruction of Wittgenstein's view of (cognitive) meaning as given by the application of semantic-cognitive rules in order to better distinguish the most relevant forms of meaning-rules and their functions. This insertion requires strong corrections in Frege's own framework. Even if the results are complex and could only be sketched here, they nonetheless seem to me clearly more auspicious than Frege's own original views.

APPENDIX TO CHAPTER IV

FREGE, RUSSELL, AND THE PUZZLES OF REFERENCE

Too much perfection is a mistake.
—*Alexandro Jodorowski*

Bertrand Russell conceived his theory of descriptions as a way to solve so-called puzzles of reference. Frege's theory of sense suggests a very different way to solve the same puzzles. While these two alternative solutions are usually assumed to be irreconcilable, each of them has its own appeal. Considering this, my proposal is that the best way to deal with this contrast is not by means of dispute, but by means of reconciliation. I will show that we can reach this reconciliation by salvaging the truth in each solution and discarding the falsity, justifying in this way their resilient appeal. More specifically, I will proceed first by removing the metaphysical load from each of these views and then by showing that with the help of appropriate adjustments, a bridge between Russell's and Frege's solutions will be built making them fully compatible, since they are only two different ways of saying the same thing.

1. Russell's solutions to puzzles of reference

I will first present Russell's four puzzles and his solutions by means of his theory of descriptions (Russell 1905: 479-493; 1919, Ch. XVI).

(i) *Reference to the non-existent:* Consider first a statement whose grammatical subject does not refer to anything, 'The present King of France is bald.' How can we attribute baldness to someone who *does not exist*?

Russell's response is that this problem only arises when we understand a definite description like (1) 'the present King of France' as a referential expression functioning as a proper name. But we can easily show that it actually does not function in this way. Letting K abbreviate '…is a present

King of France' and letting B abbreviate the predicate '…is bald,' the theory of descriptions allows us to symbolize the 'The present King of France is bald' as (2) '($\exists x$) [(Kx & (y) (K$y \rightarrow y = x$)) & Bx].' Or, to use an intuitively clearer formulation, we get the following *false* sentence:

(3) There is at least one x and at most one x, such that x is a present King of France and x is bald.

In these last two formulations, one thing is clear: there is no baldness predicated of a present King of France. When the definite description 'the present King of France' is replaced by quantified predicates, it becomes clear that we do not need to assume the existence of any present King of France to whom we should apply the predicate baldness. Moreover, since the first statement of the conjunction is false, the whole statement must be false.

(ii) *Negative Existential:* The second puzzle concerns the apparent impossibility of denying the existence of an object when the expression that denies the existence is about the same object. The problem assumes a striking form when we consider the following two statements:

1. The present King of France does not exist.
2. Statement (1) is about the present King of France.

Both statements seem to be true. However, they are mutually inconsistent. If statement (2) is true because it claims that statement (1) is about the present King of France, (1) must be false and vice versa.

Russell solves the riddle by suggesting that statement (2) is false. In order to show this, he interprets the negation in statement (1) as possessing wide scope in relation to the definite description. The analyzed form of statement (1) is (3) ~($\exists x$) [Kx & (y) (K$y \rightarrow y = x$)]; more intuitively:

4. It is not the case that there is at least one x and at most one x, such that x is a present King of France.

This is a *true* sentence since it is the negation of a false conjunction. However, it does not commit us to the existence of the present King of France, since it only commits us to denying the existence of at least one and at most one thing that has the property of being a present King of France.

(iii) *Identity Statements:* A third puzzle is the Fregean paradox of identity. Consider the statement: (1) 'The author of *Waverley* is Scott.' It contains two referential expressions, both referring to the same object. But if this is so, then statement (1) should be tautological, stating the same thing as (2) 'Scott is Scott.' However, we definitely know that (1) is a contingent and informative statement and not a tautology. Why?

Once more, Russell's solution is to make the definite description disappear. Letting *s* abbreviate the name 'Scott,' *w* abbreviate '*Waverley*' and A abbreviate the two-place predicate '…is the author of…,' we can paraphrase the identity statement (1) as (3) '$(\exists x)$ [Axw & (y) (A$yw \rightarrow y = x$) & $(x = s)$].' More intuitively:

> 4. There is precisely one *x* who is the author of *Waverley,* and this *x* is Scott.

From these last two formulations, it is clear that (1) is an informative statement since there is no doubt that its analyzed form (4) is an informative statement, very different from (2).

(iv) *Intentional context:* A final riddle that the theory of descriptions is expected to solve is that of inter-substitutability in statements of *propositional attitudes.* These statements express relational states connecting a mental attitude expressed by verbs like 'believe,' 'desire,' 'hope,' 'think,' 'want'… to what I here prefer to call a thought-content (e-thought, proposition). Consider, for instance, the two following statements:

(1) George IV believes that Scott is Scott.
(2) George IV believes that the author of *Waverley* is Scott.

Statement (1) is true since George IV was certainly able to apply the principle of identity to a proper name. However, since the name 'Scott' and the description 'the author of *Waverley*' refer to the same person, it seems that here we can apply the principle of identity substitution. It seems that we can replace the first occurrence of the word 'Scott' in statement (1) with the description 'the author of *Waverley*,' obtaining statement (2), 'George IV believes that the author of *Waverley* is Scott,' so that (2) will preserve the truth-value true. However, this does not happen: it may well be that statement (2) is false simply because George IV does not know that the author of *Waverley* is Scott, despite the obvious truth of (1). Why is this so?

In order to answer such objections Russell uses his theory of descriptions, paraphrasing (at least in relevant cases) (2) with statement (3) 'George IV believes that ∃*x* [A*xw* & (*y*) (A*yw* → y= x) & (x = s)].' More intuitively, we can express (3) as:

4. George IV believes that there is at least one *x* and at most one *x*, such that *x* is the author of *Waverley* and that this *x* is Scott.

Certainly, this is an informative belief, clearly distinct from the tautological belief that Scott is Scott. This is why George IV can believe in (1) and disbelieve (2).

2. Fregean solutions to the same puzzles

Frege has explicit answers to the last two puzzles of reference. As for the first two, we can only presume how should be the Fregean solutions.

(i) *Reference to the non-existent:* Frege suggested that in a scientific language a singular term without reference could refer to an empty set. If we try to apply this suggestion to natural language, the sentence:

(1) The present King of France is bald,

should be false, since the empty set isn't bald. However, in addition to being arbitrary, this suggestion would lead to absurd conclusions, such as that the statement 'Pegasus = the present King of France' is true, since both singular terms, 'Pegasus' and 'the present King of France' refer to the same thing, namely, the empty set.

The alternative I would like to propose starts from the notion that we can say things about non-existents insofar as the corresponding empty singular terms still preserve their senses, that is, their identification rules, even if only roughly sketched. Once we have these senses-rules in mind, we are still able to say something about their objects, not as real ones, but *merely as conceivable ones*. This is the case of the present King of France, a title which has a sense-rule, allowing us to apply it only in our imagination, thinking of France today as a Kingdom like Belgium... In this way, we are still able to articulate in rehearsal the sense-rule of the predicate with the sense-rule of the singular term. This allows us to understand Frege's sentence (i) 'Odysseus, while sleeping, was set ashore in Ithaca', which has no real reference, but only an imaginable one.

According to Frege's view, the thought-content of a sentence such as (i) should have no truth-value: since if a part of a thought (Odysseus) has no reference, the thought as a whole is also devoid of reference, devoid of truth-value (1892: 32-33). P. F. Strawson influentially supported this view, considering such statements to have what some today call 'truth-value gaps' (*Cf.* Strawson 1971: 85). This view is opposed to that of Russell's theory of descriptions, according to which statements such as (i) must be *false*, as for him 'Odysseus' should be the abbreviation of a bundle of definite descriptions without reference.[1] (*Cf.* Russell 1912, Ch. 5)

As to the question of the truth-value of statements without reference, after more than half a century of disputes, it seems to me clear that the strongest arguments favor Russell. First, it seems definitional that a proposition (e-thought-content-rule) is the kind of thing that for intrinsic reasons given by its function of *saying something that has a minimal amount of informative usefulness* must be able to have a known or at least an unknown but in some way possibly known truth-value. Second, although one might doubt that the statement 'The present King of France is wise' (Strawson) is false, just a little reflection will show that it is more reasonable to view it as false. Consider, first, examples of statements in which the singular term is also empty, but which have predicates that have *more weight* – defining 'weight' as *the power of semantically attract our attention* – either because they have a more complex semantic structure or because they are particularly relevant or curious or puzzling. Some examples:

1. I saw the present King of France strolling on the beach last week.
2. The present King of France has forbidden tourists to visit the Palace of Versailles.
3. Yesterday the present King of France was inebriated and therefore unable to perform his official duties.
4. The present King of France visited me this afternoon and we had the opportunity to discuss the EU's inability to solve European problems.
5. The present King of France is sitting on that chair.

[1] Saul Kripke has denied this, suggesting that Russell and Frege appealed to a simplified model of descriptivism with only one definite description, while the bundle theory arose later. But we need only read with attention Chapter 5 of Russell's *The Problems of Philosophy* (1912) and Frege's remarks (1882, 1918) to see that both were well aware that proper names abbreviate complex sets of descriptions.

These statements are all intuitively perceived as false, and it seems that the reason lies in the weight of the expressions complementing the descriptions: they force us to pay attention to their complex and curious predicative informational content (1 to 4) or to something that would attract great attention if it were not glaringly false (5). We see them as false because we pay attention to the non-applicability of the predicate.

Moreover, when we say 'The present King of France does *not* exist' (the denial of the presupposition), this statement is obviously true. However, statements like this should lack truth-value according to a Strawsonian presuppositional analysis.[2]

Additional evidence for this point is the following statement considered by Stephen Neale:[3]

6. The present King of France isn't wise, because there is no present King of France.

Statement (6) seems intuitively true. But (6) could not be true if the statement 'The present King of France is wise' were *not* really false. If it had no truth-value, all of statement (6) would also be devoid of truth-value.

As some have seen (Russell 1957, III; Sainsbury 1979: 118; Blackburn 1984: 309-10), the reason why the statement 'The present King of France is wise,' chosen by Strawson, appears to lack truth-value is only a *pragmatic one*. This reason can be explained as follows. First, we normally regard a statement as false because its predicative expression does not apply while we assume that the singular term applies; for instance, the statement 'Bertrand Russell was bald' is obviously false since this is a standard case of a predicate that does not apply to its subject. This is the expected case. However, we are not used to considering the truth-value of singular statements when the singular term has no reference, since these statements only rarely appear in our language for the simple reason that it is pointless to ascribe properties to something that does not exist! This is why we hesitate to say Strawson's statement 'The present King of France is wise' is false; our first reaction is to see it as a misunderstanding if not a statement devoid of sense or pointless. However, strictly speaking,

---

[2] In his book on logic, Strawson suggested that statements without a reference like 'The present King of France is wise' have no truth-value, because in order to have truth-value such statements must assume the truth of the *presupposed* statement 'The present King of France exists.' (1952: 185)

[3] In my view, in his classical work *Descriptions*, Stephen Neale settled the case in favor of Russell's analysis (1990: 26-28).

the statement is false. Or, more weakly expressed, in this case the language-in-use has nothing to tell. And we can suggest that Russell's formal analysis exposes a universal deep layer of our natural language that sometimes seems to us artificial in the same way as the material implication exposes a universal deep layer of our natural language that often seems artificial only because it is superposed by other layers in almost all linguistically effective practical uses. Consequently, Strawson's example provides no argument against the much stronger reasonableness of the decision to generalize, treating *all* statements with void singular term in the same way, namely, as false.

Moreover, statements that put weight on the predicative expression or on what is said complementarily to the definite description, like (1), (2), (3), (4) and (5), are seen by us as patently false. Why? Not because they belong to a different category, as some would like to believe. Their falsity is clear to us because of their predicative weight. They motivate us to pay attention to their predicative or relational expressions as being clearly inapplicable, in this way satisfying our usual criterion of falsity for singular statements. However, the ultimate cause of this inapplicability is still the same as in Strawson's examples: there is no object for them to be applied to in order to make the whole statement possibly true. By contrast, statements like:

7. The present King of France is slipping.
8. The present King of France is a dunce.
9. The present King of France is a human being.

do not seem to have any truth-value. Why? Because their predicates have little semantic weight. For this reason, we focus our attention on the void subject, and since we are not used to extracting falsity from a statement when the predicate does not apply because its singular term lacks reference, we tend to see the whole statement as lacking truth-value and being devoid of sense. However, we can say that they are all false for the same reason, namely, that we cannot ascribe these predicates to a nothing, since predicate ascription is also a usual pragmatic criterion for truth attribution.

Furthermore, consider statements that in a fictional context are undoubtedly true, such as:

10. Santa Claus has a white beard.

If understood as a statement about a fictional realm (10) is obviously true. But if understood as a statement about the real world, (10) would be a statement like (1): a statement that seems to have no truth-value though it must be false. And with good reason it shows its falsity when we make a statement with a weightier predication like:

11. I trimmed Santa Claus's white beard last Christmas.

It is false because it suggests that Santa Claus is a man of flesh and blood belonging to our real world. Since this man doesn't exist, the predicate cannot apply.

Finally, it is worth noting that we can possibly construct verifiability rules for these statements, which also shows that they are meaningful, expressing e-thought-rules. One can consider ways to verify that there is no bald or wise present King of France, that there is no real Santa Claus whose beard someone trimmed last Christmas, etc. All the given statements can be directly or indirectly falsified by the absence of independent external criteria for the satisfaction of their verifiability rules.

(ii) *Negative Existential*: It is not so easy to give a Fregean explanation for the enigma of negative existentials. However, consider the following statement:

(1) The present King of France does not exist.

It is true that 'the present King of France' is a definite description that does not refer to anything. But here as well the description 'The present King of France' has at least a conceptual sense, that is, a rough identification rule whose application can be at least conceived. Now, if existence is the property of effective applicability of a semantic-cognitive rule in a proper domain or context, and the identification rule expresses by the description 'the present King of France' does not apply to any object in this context, which is here inserted in the fundamental domain of real things, our conclusion is the following. The e-thought-content-rule expressed by the assertoric sentence (1) is true, since the predicate '…does not exist' simply says that the sense, mode of presentation or identification rule of 'The present King of France' isn't satisfied, that is, this rule *isn't* applicable to any object in the present domain of real things, as suggested, though it remains applicable in a conceivable, merely imaginary domain, which makes the statement sufficiently meaningful.

The same can be said for the denial that the referent of a proper name exists. If the sense of a proper name, as Frege indirectly suggested, is the abbreviation of bundles of definite descriptions, or, as I have defended, the abbreviation of a properly characterized disjunction of fundamental descriptions, then a similar strategy is applicable to negative existential statements with empty names. Take for example statements like (i) 'Vulcan does not exist', calling 'V' '…a small planet circling the sun inside the orbit of Mercury,' we can symbolize the sentence (i) as $\sim\exists x\ [Vx\ \&\ (y)\ (Vy \rightarrow y = x)]$. What sentence (i) means is that the conceptual sense expressed by the fundamental descriptions abbreviated by the name of the small planet 'Vulcan' has *no* effective application in its proper domain, that its identification rule isn't satisfied by any real object, which is true.

(iii) *Identity Statements:* The riddle of identity between descriptions can be exemplified by the most discussed sentence of analytic philosophy:

(i) The morning star is the evening star.

For Frege this identity sentence is informative because the descriptions 'the morning star' and 'the evening star' express different senses or modes of presentation of the same object, the first as the brightest celestial body that appears to us at dawn, and the second as the brightest celestial body that appears to us in the evening…

As already seen (Ch. IV, sec. 27), particularly concerning proper names, due to their semantic flexibility, a double answer could be given depending on *different contextual emphases*. To make it easier, suppose that we have the proper names 'Phosphorus' (Morning Star) and 'Hesperus' (Evening Star) building the sentence (ii) 'Phosphorus is Hesperus.' There are two main ways of understanding this sentence, depending on which semantic element we are emphasizing in accordance with the context:

> *Immediate-derived Emphasis*: In this case, the senses, the modes of presentation for Phosphorus and Hesperus as their separate identification rules, are emphasized, Phosphorus being understood as the last star to disappear at dawn and Hesperus as the first star to appear in the evening... Here the whole mode of presentation of Venus, which contains both visible stars and is responsible for their identity, is left in the background, being only the resulting datum of an identity that we expect to preserve. In this case, the statement is seen as expressing a derived *contingent a posteriori* thought, emphasizing the

difference as opposed to the identity, being this identity informative, since it still informs us in an implicit supplementary way that these two different senses or identification rules have the same ultimate reference. The derived statement refers first to the apparent *sub-fact* that Phosphorus isn't Hesperus and only secondarily lets us infer the further grounding fact of Venus' self-identity. Its emphasized modal form can be read as ◊ (a = b). This is how Frege saw the identity.

Mediated-basal emphasis: In this case, with both names we emphasize that we mean Venus, attaching to both terms the same fundamental localizing astronomical description (say, the second planet of the solar system, etc.) that forms an accepted identification rule that has a variety of guises, of ramifications as modes of presentation, under the assumption of our current astronomical views. Here the descriptions of Venus' appearances to us play only the role of irrelevant auxiliary descriptions. Because of this, the sentence 'Phosphorus is Hesperus' is here seen as an uninformative analytic identity sentence – a *necessary a priori* sentence – even if it has different fringes of meaning depending on the different auxiliary descriptions related to different usual modes of presentation. In this case the assertoric sentence has as its most proper sense the basal thought referring to the *grounding fact* of Venus' self-identity, being expressed by the sentence 'Venus [in full] = Venus [in full],' from which we may derive 'Phosphorus [-Venus] = Hesperus [-Venus],' or 'Phosphorus = Hesperus.' These statements are *necessary a priori*, emphasizing the identity in the difference. Their emphasized modal form can be rendered as □ (a = b). This is how Saul Kripke chose to see the identity.

As was noted in the last chapter, Kripke's *necessary a posteriori* identity between proper names is the result of a confusion of the necessity of the mediated-basal emphasis with the contingency of the immediate-derived emphasis.

(iv) *Intentional contexts:* As for the enigma of intentional contexts, Frege suggests that in statements of propositional attitudes, the subordinate sentence does not have its usual reference – its truth-value – but rather an *indirect reference*, which is *its sense*. Thus, in saying (1) 'George IV believes that Scott is Scott,' the reference of the subordinate sentence 'Scott is Scott' isn't its truth-value or a corresponding fact, but simply *the thought* expressed by this sentence. And in saying (2) 'George IV believes that the author of *Waverley* is Scott,' the subordinate sentence 'the author of *Waverley* is Scott' also refers to a thought. Since the references of

'Scott is Scott' and 'the author of *Waverley* is Scott' are different, the sentences (1) and (2) cannot be interchangeable *salva veritate*.

I do not wish to discuss here the objections of detail that could be made to Russell's and Frege's solutions. I want to mention only the general objection made to Fregean-kind solutions for riddles of reference, according to which they induce us to accept some kind of Platonism of senses and thoughts, unlike Russell's ontologically more economical solutions. Against this, the last chapter made clear that we can preserve objectivity of sense as something interpersonally accessible without any recourse to ontological realism. All we need is to understand senses as embodied semantic-cognitive rules developed as interpersonally corrigible rules or conventions or as their derived adequate associations.

3. Reviewing Fregean assumptions

Who is right? Russell or Frege? As I noted at the outset, my hypothesis is that it is not a matter of choosing between two views. The fact that we have achieved no consensus regarding the right theory reinforces my suspicion that both theories have some truth. This is why I suppose that each of them has insightful content mixed with very implausible metaphysical assumptions, and that these implausible assumptions are what make them appear irreconcilable. Thus, in the course of this Appendix I will reconstruct these theories by eschewing their metaphysical assumptions and filling the resulting gaps with more plausible views.

Let's start with Frege. We have already seen that we can eliminate the anachronistic ontological realism of sense if we replace it *with any* psychological instantiation of a semantic-cognitive rule qualitatively identical to the one with which we are associating the expression. Repeating what has already been proposed in our reading of Ernst Tugendhat in the introductory chapter, it is perfectly plausible to identify what Frege called the senses of singular predicative sentences in terms of semantic rules, so that: (i) the sense of a nominative expression (the mode of presentation of the object) is the same as the *identification rule* (*Identifikationsregel*) of a singular term, whose criteria of application are adequate configurations of identifying tropes of the object; (ii) The sense of a predicative expression (as its conceptual content) is the same as its *ascription or application rule* (*Verwendungsregel*), whose criteria of application are tropes dependently associated with the object; the sense of a singular declarative sentence (its e-thought or thought-content) is the

same as its *verifiability rule* (*Verifikationsregel*) associating (i) and (ii). (See Tugendhat 1976: 262; Tugendhat & Wolf 1983, Ch. 13) To this, we have added that verifiability rules demand criteria of application which are their possible truth-makers, which can often be better identified (differing from Frege) with the *sub-fact* referred to by the statement, this sub-fact remitting to a *grounding fact* as aspects of the latter.

A second point is to reject some of Frege's odd ideas concerning reference, like those of an unsaturated concept as the reference of a predicate and of truth-value as the reference of a sentence, as I argued in the last chapter. It is much more plausible to see the concept in a natural way as the sense of a predicative expression – a conventionally grounded rule – and the reference of a sentence not as a truth-value, but simply as a fact.

A further thing we did in the last chapter was to paraphrase the Fregean concept of existence. For Frege existence was the property of a concept of being satisfied by at least one object. For us existence is the property of a possible conceptual sense – of a possible semantic-cognitive rule – of being effectively (and not merely putatively) applicable to at least one referent belonging to a chosen domain or context (usually the most proper domain or context) during some period of time (the period in which the object is said to exist). Thus, to know that a referent exists is to know that its conceptual rule, if it exists, is effectively and continuously applicable in its most proper domain or context in the time during which the referent (a tropical property, an object, a fact) can be said to exist. Moreover, as we have seen, this does not deprive existence of objectivity, because if the effective applicability of a conceptual rule is a tropical property of the rule, it is also a higher-order tropical property of the referent, which is that of having its conceptual rule effectively applicable to it – even if this rule was never applied and even if it does not exist as an actuality but merely as a possibility (a dispositional or possibility-trope)! – if the right conditions were given, the rule would exist and be definitely applicable. This is a minimal condition allowing us to envisage an object as really existing in the outside world.

This result can be conceded for each of the rules (senses) already suggested in Tugendhat's analysis of singular statements: (i) The existence of an *object* (made up of a certain relatively independent compresent cluster of tropes) is the same as the effective applicability of its proper identification rule to itself. (ii) The existence of a property-trope – differing from the object to which it is attached by a relative dependence – is the same as the effective applicability of its ascription rule to itself. (iii) By symmetry with cases (i) and (ii), the existence of a *fact* in the world

(minimally an arrangement of an independent cluster of compresent tropes and a dependent property-trope) is the same as the effective applicability of the verifiability rule constitutive of the e-thoughts to the verifier (truth-maker) of this fact. Since the verifiability rule is the real Fregean thought, the existence of the fact is also the effective applicability of its thought, expressible by an assertoric sentence. Existence here, as you remember, is also called 'truth' in the derivative sense of the *reality* of a fact.[4]

Finally, even in the context of Fregean theory, I want to treat sentences without a reference as ultimately *false* and not as simply devoid of truth-value, as Frege suggested. After all, the reason Frege believed that sentences with components that lack reference are devoid of truth-value lies in his insistence on the indefensible idea that a sentence's reference should be its truth-value. However, at this point we are already certain that a sentence's reference is a fact. Therefore, the absence of such a fact just leads us to the falsity of the whole sentence, as we have shown in our discussion of the Fregean solution to the question of the reference of non-existents. This heavily corrected version of Frege's view is already close to the position held by Russell, who regarded sentences with empty attributive definite descriptions as false.

4. Reviewing Russellian Assumptions

Now it is time to review the assumptions of Russell's theory of descriptions. A first step is to rule out (i): his thesis according to which:

> Definite descriptions and even our usual names (which for him were clusters of descriptions) are not to be viewed as referential terms, but rather as *incomplete symbols*. (*Cf.* Russell 1994)

[4] Certainly, all three cases can be expressed using formal devices in which referential terms are transformed into predicative expressions. Thus, consider the existence of what is *predicated* in the statement 'Marsupials exist': symbolizing '…is a marsupial' as M, we have '(∃x) (Mx).' Consider now the *definite description* in the statement 'The Morning Star exists': symbolizing the predicate '… is a morning star' as M, we have '∃x [Mx & (y) (My → y = x)].' For the *proper name* in the statement 'Socrates exists,' abbreviating the complex descriptive content that the name contains with the predicate 'socratizes' and symbolizing this last predicate as 'S', we have (∃x) [Sx & (y) (Sy → y = x)]. Finally, consider the *statement* 'Socrates is wise': symbolizing '…is wise' by W, we have (∃x) [Sx & (y) (Sy → y = x) & Wx].

This Russellian thesis flies in the face of our most fundamental natural language intuitions. For what could better exemplify a referential expression than a proper name or even a definite description? One could even say that our usual proper names, definite descriptions, and indexicals, are *patterns* of singular referential terms *whose definitional function is to select precisely one object, indicating which it is among all other objects of a certain domain*.[5] The attempt to change this is to distort natural language in a way that only serves to spread confusion. Thus, without denying that definite descriptions are incomplete symbols, I will maintain that definite descriptions are patterns of referential terms.

Russell's intention with his logical atomism and semantic referentialism was to eschew the supposed referential and semantic role of definite descriptions with the ultimate goal of replacing natural language referential expressions with what he called *logically proper names* – the only truly referential expressions. However, as we have already seen earlier in this book (Ch. III, sec. 3), this doctrine is hopeless, and his semantic referentialism indefensible (*Cf.* Tugendhat 1976: 437; also Kripke 2013, Ch. 1).

Once we reject Russell's atomistic doctrine of logically proper names, there is no reason to deny that ordinary names and definite descriptions are referential terms. Even when a definite description is analyzed in the form of a conjunction of quantified predicative expressions, as Russell does, it can continue to do the same referential work of a singular term, since it is assumed that the definite description is able to pick out a single object and distinguish it from all other objects of a given domain. This is all that is required for an expression to be a singular term.

We must also reject a second assumption made by Russell, namely, his strange suggestion that (ii) *definite descriptions do not have any meaning in themselves*. As he wrote:

> I advocate that a denoting phrase is essentially *part* of a statement, and does not, like most single words, have any significance on its own account.' (1994: 51)

This assumption makes sense within the semantic referentialism of Russell's logical atomism: since for him descriptions aren't referential expressions and reference is the source of meaning, it is justified to say that they aren't intrinsically meaningful. But even if you complete them by

[5] As Ernst Tugendhat pointed out, in opposition to Donald Davidson, to refer to one object is not only to coordinate the name with it but 'to distinguish it from all the others belonging to a certain domain.' (Tugendhat & Wolf 1983: 153)

constructing meaningful statements like 'The man who wrote "On Denoting" was a philosopher,' it seems impossible to explain why the addition of a new predicate produces a meaningful statement. Assumption (ii) only reaffirms the incoherence of Russell's semantic referentialism. One cannot reasonably doubt that definite descriptions have meanings in themselves or that they are referential expressions.

Now, once we reject Russell's semantic referentialism and admit that we usually make our references by means of semantic-cognitive rules, one thing is clear: the Russellian requirement of applying a predicate to a single object with such-and-such characteristics already constructs something at least close to an identification rule with a complete sense allowing us to refer to something unique.[6]

5. Building a bridge between both views

Once in possession of a metaphysically unspoiled understanding of Frege's and Russell's analysis – one that strips them of their implausible speculative wrappers – we are ready to take the final step. We need to use the semantic-cognitive rules constitutive of senses, together with the concept of existence as the effective application of these rules, in order to build a bridge allowing us to travel from Fregean solutions for riddles of reference to Russellian ones and vice versa. In this way, I will demonstrate that their answers to puzzles of reference are in essence inter-translatable and therefore reconcilable. Here is how this can be done:

(i) *Reference to non-existents.* As we have seen, the most reasonable answer to the Fregean problem of how to give meanings to statements referring to non-existent objects is that we can at least conceive how we can supplement the dependent (unsaturated) sense of a predicative expression with the independent (saturated) sense of a singular term, thus constituting the complete content of a thought. This is what allows us to think of the present King of France as bald or wise… without having to admit his actual existence.

A better understanding emerges when we translate Fregean senses in terms of semantic-cognitive rules. In this case – following Tugendhat – we

[6] I will leave aside all the complexities related to 'non-Russellian' definite descriptions like 'the round table in this room' (indexical use), 'the man drinking a martini over there' (referential use), 'the White Anglo-Saxon Protestant' (general use), 'the reason why I like beans' (justifying use)... All they do here is to divert us from our intended point, creating specious distractions.

normally say that the true ascription rule of the predicate always applies to its usual reference as a consequence of the application of the identification rule. Returning to an example considered in the introduction of this book: Seeing the Earth from outside the Earth's atmosphere for the first time, Russian cosmonaut Yuri Gagarin remarked: 'The Earth is blue.' But in order to formulate this thought, he first had to identify something outside his space capsule, an object, the planet Earth. Only *by means* of this identification could he apply the predicate '…is blue' to the trope of blue belonging to the object he had visually located. We see that the rule for the application of the predicate '…is blue' needs to be first, say, *driven* by the selective application of the identification rule in order to find the object called 'Earth,' only then being able to be applied in the identification of the particularized property-trope of this object of being blue.[7]

Let us now consider the case of empty singular terms, the alleged reference to non-existents, as found in the sentence 'Vulcan is red.' According to the calculations of the astronomer Le Verrier, Vulcan should be a small planet located in an orbit approximately 21 million kilometers from the Sun… Now, this is the Fregean *sense* of this name, the mode of presentation of its reference, the identification rule for the planet Vulcan. However, since we are now certain that the planet Vulcan does not exist, we know that the name's reference is empty, that its identification rule is inapplicable. As a result, the effective application of the ascription rule of the predicate '…is red' is also impossible. As the identification rule of the singular term doesn't apply to any expected object, an application of this rule cannot be made, remaining non-satisfied by any actually given cluster of tropes. Thus, the predicate cannot be applied, making the sentence false (*pace* Frege and Strawson).

As noted above, we do not need complex metaphysical theories to explain what happens in this case. The right explanation appeals to our *capacity for imagination*. We are at least in some measure always able to *conceive* what it would be like to apply both rules in association, even if

[7] Surely, Gagarin could also say 'The blue thing out there is the Earth.' But then he would use 'The blue thing out there' as a singular term and the 'is' ('…is the same as…') as expression of the relational ascription rule applicable only after the application of the two identification rules. On the other hand, if the statements were 'The Earth is red' or 'Out there is red,' they would be false because the object/place located by means of the identification rule would not have the property-trope able to satisfy the ascription rule of the predicate '…is red' (*Cf.* Ch. I, sec. 2)

we cannot find a way to apply them to the real world.[8] To use a Wittgensteinian expression, we are able to conceive the application of a statement like 'Vulcan is red' as a *possible state of affairs* (1984a, 3.02). It is only to the extent that we are able to *conceive* the possibility of applying both rules in the constitution of a verifiability rule that we can understand the cognitive meaning of the statement. When we do this, we realize that the proper name is empty and that the e-thought-rule (cognitive meaning, verifiability rule) that contains it has no effective application to a real fact in the world. This is why the statement 'The present King of France is wise' is already able to express a complete sense as an e-thought-rule. We are capable of conceiving the two rules used in association in order to form the verifiability e-thought rule, the sense of the statement, imaginatively applicable in our minds to a possible fact, but without effective application in its most proper domain as a real fact in the world; and this impossibility makes this e-thought-rule false.

As to the question of how it is possible to assign baldness or wisdom to a non-existent person, the answer is now clear: we are capable, at least in some measure, of conceiving the application of semantic-cognitive rules and their adequate associations, and by doing this we give meaning to the terms and the sentence as a whole. Mentally we are able to make a fictitious predication, even if only to a limited degree, without endowing it with a proper assertoric and judicative force.

Now, in light of this reconstruction, it is easier to make the theory of sense agree with the theory of descriptions. We can paraphrase the definite description of the statement 'The present King of France exists' in a Russellian way as:

1. There is at least one *x* and at most one *x*, such that *x* is a present King of France.

And we can say that what is expressed here (disregarding the attribution of existence) is a somewhat abbreviated formulation of the Fregean sense of the same identification rule for the present King of France, which is seen as having two components:

(i) the condition of uniqueness,

[8] I use *'conceive'* and *'imagine'* as equivalent verbs thought with different emphases. In a broad sense, not all imagination is imagistic. We can speak, for instance, of 'mathematical imagination'.

(ii) the ascription rule for the predicative expression '…is a present King of France.'

Together (i) and (ii) constitute a kind of identification rule, because they give us the possibility to distinguish at least one and at most one object by means of criterial properties derived from the predicate, such as the supposed existence of a hereditary head of state governing France today.

The non-existence of the present King of France corresponds to the lack of effective applicability of the identification rule roughly expressed by the conjunction of (i) and (ii) and, therefore, to the lack of a reference. As for the predicate '…is wise' in the sentence 'The present King of France is wise,' its ascription rule also does not apply, since no one has the property of being the present King of France to whom the rule *could* apply. Anyway, this predicate still expresses an ascription rule as something only conceivably applicable, a conceivable Fregean sense understood as a mode of presentation.

Pulling the threads together, with the statement 'There is at least one *x* and at most one *x*, such that *x* is a present King of France, and *x* is wise' we do nothing more than try to apply the same verifiability rule expressed by the statement 'The present King of France is wise.' That is, we realize that the identification rule cannot find a bearer and that consequently, the ascription rule is also inapplicable, the same being the case with their adequate associations in the form of a verifiability rule. In this way, analyzing the case of reference of non-existents, we are already able to see how we can exchange a 'Fregean' explanation for a 'Russellian' explanation and vice versa.

(ii) *Negative Existentials.* In the last chapter (despite Frege's view) we identified the concept with the sense of a predicative expression. This also means that to say 'The present King of France does not exist' becomes the same thing as saying that the sense of 'the present King of France' does not determine its reference.

How would we express this using semantic-cognitive rules in place of the sense? Well, we would again say that the sense or meaning expressed by a singular term like 'the present King of France' consists in the identification rule of this definite description in its only conceivable application. We know this because we know we can at least to some extent imagine how we would apply this definite description. But we cannot gain any awareness of the effective applicability of this rule, that is, we cannot say that the object that should be referred to by this definite description

exists, since we know that this rule cannot be definitely applied in its most proper context.

Finally, we come to the corresponding 'Russellian' analysis. A description like 'the present King of France' is here transformed into

1. at least one *x* and at most one *x* is such that *x* is a present King of France.

Here again, what we have is an identification rule for a particular object, which is composed of two sub-rules:

(i) a rule demanding unity,
(ii) a rule of application of the predicate '...is a present King of France.'

Now, to say, 'The present King of France does *not* exist,' is to say:

It is not the case that there is at least one *x* and at most one *x* such that *x* is a present King of France.

But this is the same thing as to say that the identification rule roughly composed of conditions (i) and (ii) is *not* effectively applicable. What is the difference between this rule and the Fregean sense of the description? The answer is again that the 'Russellian' analysis only decomposes the identification rule of the definite description 'the present King of France' into two rules: a unity rule and a rule of application for the predicate. Saying that *the* present king of France does *not* exist is to say that the ascription rule of the predicate '…is a present King of France' does *not* effectively apply in its proper context, in this case, because it does not fulfill the implicit existential condition. Once more, the 'Russellian' and 'Fregean' analyses of negative existentials reveal themselves as two different ways to say almost the same thing.

(iii) *Identity.* Consider now identity sentences like 'The Morning Star is the Evening Star.' How can this sentence be informative, if the two descriptions refer to the same object? Frege's reply is that despite the fact that these descriptions refer to the same object, they express different modes of presentation of this object, being therefore informative.

Paraphrasing the concept of meaning in terms of a semantic-cognitive rule, what a Fregean semantics suggests is that the sentence above is informative because it tells us that we identify the same object using (a)

two different identification rules, or (b) two branches of the same identification rule. These rules or branches are respectively a rule for the identification of the last star to disappear at dawn and a rule for the identification of the first star to appear in the evening. These rules call for different criterial settings, emphasizing the apparent sub-fact of the difference between the Morning Star and the Evening Star. That in the end they refer to the same object is – in the context considered by Frege – a *further* piece of information, a complementary identification rule for the planet Venus. If we add this last piece of information in order to build a *unifying rule* requiring assumptions about our astronomical knowledge, we have a conditioned *necessary a priori* e-thought-rule. Otherwise, the e-thought-rule is seen as *contingent a posteriori* (See section 2 above; also Ch. IV, sec. 26).

In Russellian terms, letting M abbreviate the predicate '…is a morning star' and E abbreviate the predicate '…is an evening star,' the identity sentence can be symbolized as:

(1) $\exists x\,[(Mx \,\&\, Ex) \,\&\, (y)\,(My \rightarrow y = x) \,\&\, (z)\,(Ez \rightarrow z = x)]$.

In other words:

(2) There is precisely one *x* that is the morning star and *this same x* is also the evening star.

In this case, what we are doing with the identity sentence is (i) making a conjunction of two different ascription rules of predicates and adding to it the condition (ii) that they both apply to one and the same object. Thus, the 'Russellian' analysis only assures us that the identification rule constituted by '$\exists x\,[Mx \,\&\, (y)\,(My \rightarrow y = x)]$' applies to the same object that the identification rule constituted by '$\exists x\,[Ex \,\&\, (z)\,(Ez \rightarrow z = x)]$' applies to, since by transitivity, if $y = x$ and $x = z$, then $y = z$. But this is already near to the claim that we have two different identification rules, two different Fregean modes of presentation, further known as having the same object. Again, the two analyses turn out to be largely interchangeable.[9]

(iv) *Intentional Contexts.* Finally, consider expressions of propositional attitudes such as:

[9] It is not our topic here, but it is worth noting that in any case the identity can be seen as *necessary a priori* insofar as we take for granted our astronomical knowledge. In this case, the identity is a priori and conditionally (hypothetically) necessary, and both identification rules are aspects of a single, more complex identification rule of Venus.

(1) George IV believes that Scott is Scott.

And

(2) George IV believes that the author of *Waverley* is Scott.

Why doesn't the truth of (1) guarantee the truth of (2), if both subordinate clauses are identity sentences about the same person?

As we have noted, for Frege the answer is that in such cases a subordinate clause does not have its usual reference, which for him is its truth-value. Subordinate clauses, he holds, refer to the *thoughts* expressed by them, and the thoughts expressed by them in (1) and (2) are different. Hence, the truth-value of the whole sentence that expresses a propositional attitude cannot depend from the truth-value of the subordinate clause, which makes inter-substitution *salva veritate* impossible.

Since we reject Frege's artificial idea that a sentence's normal reference should be its truth-value, we must first reformulate his solution. For us, an isolated statement such as 'The author of Waverley is Scott' has as its immediate reference the aspectually given sub-fact represented by the identification of the modes of presentation of the singular terms flanking the identity relation. This sub-fact can be represented by the statement 'Being the author of Waverley ≠ being Scott,' while the mediated reference, the grounding fact, can be represented by the statement 'Scott = Scott' (The underscore '_' signals that I am speaking about facts). As already explained, both facts are complex tropical arrangements.

Now, what *fact* is represented in the case of propositional attitudes? First, we can preserve Frege's idea that in utterances of propositional attitudes the reference of the subordinate sentence is its *sense*, for us an e-thought-content that is ultimately a *mental fact*. But there is more to the matter. This mental fact is part of the whole fact represented by a propositional attitude, which has the form *aAp*, in which *a* abbreviates the relevant descriptions identifying the person who has the attitude, *p* is the subordinate sentence referring to *a*'s *e-thought-content*, and *A* abbreviates the attitudinal verb applied by *a* to *p*, which can be one of belief, knowledge, desire, etc. Hence, the reference of 'Henry IV believes the author of *Waverley* is Scott' is no typical fact in the external world. It is a fact consisting in the *psychological belief of the real Henry IV* that the author of *Waverley* is Scott. In other words, a propositional attitude conventionally refers to an *essentially mental fact*: *the (mental) attitude of*

a (partly non-mental) speaker (a person[10]*) concerning a certain (mental) thought-content* that we can symbolize as *a<u>Ap</u>*. Here *p* refers to a thought-content (dispositional or not) in the mind of person *a*, such that it no longer refers to any fact in the external world that could possibly match *p,* making it true. Here, if ⊢*aAp* affirms the essentially mental fact that *a<u>Ap</u>,* then the statement is true, otherwise it is false; and while as a person *<u>a</u>* should be a cluster of compresent (physical and mental) tropes in the world, *<u>Ap</u>* distinguishes itself by being a mental relational tropical complex appropriately linking person *a* with a factual arrangement of her own mental tropes. In other words, what matters in statements of propositional attitudes is a certain relationship between the contents of the main clause (usually expressing the speaker's dispositional mood or mental act) and the e-thought-content expressed by the subordinate clause. And it must be that the truth of a sentence of propositional attitude depends only on the *fact* of this attitudinal relationship A to p really being in *person a's mind*, independently of the truth or falsity of the thought-content expressed by p concerning any independent fact in the real word.

We can now see more clearly why the thought expressed by the subordinate clause cannot be replaced *salva veritate* in (1) and (2): in each case *a's* mental dispositions or acts concern different factual thought-contents expressed by different subordinate clauses. Finally, it is worth noting that the person who judges these propositional attitudes is a third person or even the first person in an introspective mood, or in a later time, and there is no distinction between the senses and the facts reported when the ascription is true.[11]

Now, to paraphrase thought-contents as verifiability rules for sentences, we need only note that the verifiability rules of the sentences of (1) and (2)

---

[10] Here I understand a person in P. F. Strawson's sense as an object to which both (physical) p-predicates and (mental) m-predicates are ascribed. (1959, I, Ch. 3)

[11] Since the reference is determined by the sense, for a Fregean there must be a second indirect sense here determining the indirect sense of the subordinate clause. But no one was able to point to this hidden indirect-indirect sense or to the regress that it might be apt to produce. We circumvent this by holding that the whole attitude described by ⊢*aAp* is first that of a third person (or the same person in a reflexive mood) concerning the essentially mental fact that *a<u>Ap</u>*. If this fact exists, *aAp* is true, otherwise not. For example: '[I am sure that] Anna believes that Goya painted the *Third of May, 1808*'. Here the fact that Anna believes that Goya painted the *Third of May, 1808*, must have an external mode of presentation *for me.* This could be because we visited the Prado Museum yesterday, which determines the reference or fact, in a case where I use an e-thought to refer to Anna's belief in her own thought-content.

are different, applying only to the essentially mental fact of the kind *aAp*, without committing us to the effective applicability of *p* to any real fact in the world. Thus, considering the sense of the proper name Scott as an identification rule, we can in many cases paraphrase (1) as:

> (1') George IV believes that the identification rule (a) (sense (a)) that he has for 'Scott' applies to the same object as the identification rule (a) (sense (a)) that he has for 'Scott.'

This tautological belief is true even if George IV knows nothing about Scott. We can paraphrase (2) as:

> (2') George IV believes that his identification rule (a) (sense (a)) for 'Scott' applies to the same object as the different identification rule (b) (sense (b)) that he has for 'the author of *Waverley*.'

The obvious argument drawn from this is the following:

1. The truth-value of the propositional attitude statements depends on the existence of the proper (essentially) mental fact that an e-thought-content *p* is the object of person *a's* attitude *A*.
2. The (essentially) mental facts represented by (1') and (2') are *different* because the e-thought-contents *p* are different.
3. Statements (1') and (2') are analyzed forms of (1) and (2).
4. Conclusion: We do not need to preserve the same truth-value in statements (1) and (2).

The two subordinate clauses cannot replace one another *salva veritate* because they have different factual e-thoughts-contents *p*'s as references and so also the two whole attitudinal statements.

Finally, consider the Russellian paraphrases. Statement (1) can be formulated as 'George IV believes that $\exists x$ [(x = s) & (y) ((y = s) → (y = x)) & (s = s)],' or simply as:

> (1'') George IV believes that there is at least one *x* and at most one *x*, such that *x* is Scott, who is the same as Scott.

And statement (2) can (in a secondary occurrence) be formulated as 'George IV believes that $\exists x$ [A*x*w & (*y*) (A*y*w → *y* = *x*)) & (*x* = s)]' or, more naturally:

(2'') George IV believes that there is at least one *x* and at most one *x*, such that *x* is the author of Waverley and *x* is Scott.

Now, as the subordinate clauses expressing George IV's beliefs (i) 'there is precisely one *x* that is Scott' and (ii) 'there is precisely one *x* that is the author of Waverley' are different, 'Scott is Scott' cannot mean the same as 'Scott is the author of *Waverley*.' The e-thought-rules expressed by (i) and (ii) are different and so also the sub-facts conceived by George IV.

It should be remarked that our version of Russellian analysis and our version of Fregean analysis have different emphases. After all, we can present the Fregean analysis in (2') for example, as:

(2''') George IV believes there is at least one *x* and at most one *x*, such that the rule of identification (a) for Scott (sense (a)) and the rule of identification (b) for the description 'the author of *Waverley*' (sense (b)) apply to *x*.

But (2') and (2''') do not differ significantly in what they say. After all, suppose we say, based on Russell, that George IV believes the rule of identification (a) that he knows for the name 'Scott' and the ascription rule (b) that he knows for the predicate '…is the author of Waverley' effectively apply to *precisely one and the same object.* This amounts to almost the same thing as to say, based on Frege, that George IV believes that the identification rule (a) (the sense (a)) he knows for the singular term 'Scott' has the *same referent* as the rule of identification (b) (the sense (b)) of the definite description 'the author of *Waverley*.' Now it is clear: also in the case of propositional attitudes, the Fregean and Russellian analyses are at least to a great extent intertranslatable.

5. Conclusion

Summarizing, we can analyze the referential function of definite descriptions in at least three ways: (a) in terms of abstract entities, as did Frege when speaking of senses, (b) in terms of semantic-cognitive criterial rules inspired by approaches like those of Tugendhat and Dummett, and (c) using resources from predicative logic, as Russell tried to do in his theory of descriptions. These are only three complementary endeavors to say the same thing.

As I have noted, the initial impression of strangeness of the proposed view comes from the acceptance of the metaphysical assumptions that permeate what Frege and Russell wrote on the issue. Against Russell's

own belief, his paraphrases of definite descriptions are nothing more than limited expressions of semantic-cognitive rules. These paraphrases make it possible to express the referential function of definite descriptions in their attributive use by means of quantified predicative expressions used in a domain that grants them a singularizing application. In this reading, they are reformulations of senses or modes of presentation that cannot be more than semantic-cognitive criterial rules. Assuming that these last rules only exist in their applications – either in imaginative psychological rehearsals or in effective cognitive instantiations concerning real entities in the world – the compatibility of the so-understood theory of descriptions with our cognitivist approach is clear.

CHAPTER V

VERIFICATIONISM REDEEMED

> There is no distinction of meaning so fine as to consist in anything but a possible difference in practice.
> —*C. S. Peirce*

> *Es ist schwer einem Kurzsichtigen einen Weg zu beschreiben. Weil man ihm nicht sagen kann: 'Schau auf dem Kirchturm dort 10 Meilen von uns und geh' in dieser Richtung.*
> [It is difficult to tell a near-sighted man how to get somewhere. Because you cannot say to him: 'Look at the Church tower ten miles away and go in that direction.']
> —*Wittgenstein*

Verificationism is now commonly viewed as a relic of philosophy as practiced in the first half of the 20th century. Although initially advocated by members of the Vienna Circle, it soon proved unable to withstand an ever expanding range of opposing arguments, which came from both within and outside of the Circle. My aim in this chapter is to show that we can achieve an understanding of verifiability that is both intuitively acceptable and resistant to the most widespread objections. In my view, the Vienna Circle failed to successfully defend verificationism because it used the wrong approach of beginning by formally clarifying the principle of verification initially proposed by Wittgenstein without paying sufficiently detailed attention to what we really do when we verify statements. When their arguments in favor of the principle were shown to be faulty, most of them, along with their offspring, unwisely concluded that the principle itself should be rejected. In my view, they were exhibiting the same reaction of the proverbial fox in Aesop's fable: unable to reach the grapes, he consoled himself by imagining they were sour...

Returning to the methodology and assumptions of the later Wittgenstein, my aim in this chapter is twofold: first to sketch a plausible version of what I call *semantic verificationism*, which consists in the proposal that the epistemic contents of declarative sentences, that is, the e-thought-contents or propositions expressed by them are constituted by

their verifiability rules; second, to confirm and better explain semantic verificationism by answering the main counter-arguments.

1. Origins of semantic verificationism

The first point to be remembered is that, contrary to a mistaken popular belief, the idea that a sentence's meaning is its method of verification didn't stem from the logical positivists. The first to propose the principle was actually Wittgenstein himself, as members of the Vienna Circle always acknowledged (*Cf.* Glock: 354). Indeed, if we review his works, we see that he formulated the principle in 1929 conversations with Waismann and referred to it repeatedly in texts over the course of the following years. Furthermore, there is no solid evidence that he abandoned the principle later, replacing it with a merely performative conception of meaning as use, as some have argued. On the contrary, there is clear evidence that from the beginning his verificationism and his subsequent thesis that meaning is a function of use seemed mutually compatible to him. After all, Wittgenstein did not hesitate to conflate the concept of meaning as verification with meaning as use and even with meaning as calculus. As he said:

> If you want to know the meaning of a sentence, ask for its verification. I stress the point that the meaning of a symbol is its place in the calculus, the way it is used.[1] (2001: 29)

It is always advisable to check what the original author of an idea really said. If we compare Wittgenstein's verificationism with the Vienna Circle's verificationism, we can see that there are some striking contrasts. A first one is that Wittgenstein's main objective with the principle always seems to have been to achieve a grammatical overview (*grammatische Übersicht*), that is, to clarify central principles of our factual language, even if this clarification could be put at the service of therapeutic goals. On the other hand, he was always against the positivistic-scientistic spirit of the Vienna Circle, which in its incipient and precocious desire to develop a purely scientific philosophy had as the strongest motivation to develop the verification principle to use it as a powerful reductionist

[1] Wittgenstein's best reader at the time, Moritz Schlick, echoes a similar view: 'Stating the meaning of a sentence amounts to stating the rules according to which the sentence is to be used, and this is the same as stating the way in which it can be verified. The meaning of a proposition is the method of its verification.' (1938: 340)

weapon, able to vanquish once and for all the fantasies of metaphysicians. Wittgenstein, for his part, didn't reject metaphysics in this way. For him, the metaphysical urge was a kind of unavoidable dialectical condition of philosophical inquiry, and the truly metaphysical mistakes have the *character of depth* (Wittgenstein 1984c sec. 111, 119). Consequently, metaphysical errors were intrinsically necessary for the practice of philosophy as a whole. As he wrote:

> The problems arising through a misinterpretation of our forms of language have the character of depth. They are deep disquietudes; their roots are as deep in us as the forms of our language and their significance is as great as the importance of our language. (1984c, sec. 111)

It was this rejection of positivistic-scientistic reductionism that gradually estranged him from the Logical Positivists.

In these aspects, Wittgenstein was much closer to that great American philosopher, C. S. Peirce. According to Peirce's *pragmatic maxim*, metaphysical deception can be avoided when we have a clearer understanding of our beliefs. This clarity can be reached by understanding how these beliefs are related to our experiences, expectations and their consequences. Moreover, the meaning of a concept-word was for Peirce inherent in the totality of its practical effects, the totality of its inferential relations with other concepts and praxis. So, for instance, a diamond, as the hardest material object, can be partially defined as something that scratches all other material objects, but cannot be scratched by any of them.

Moreover, in contrast to the positivists, Peirce aimed to extend science to metaphysics, instead of reducing metaphysics to science.[2] So, he was of the opinion that verifiability – far from being a weapon against metaphysics – should be elaborated in order to be applicable to it, since the aim of metaphysics is to say extremely general things about our empirical world. As Peirce wrote:

> But metaphysics, even bad metaphysics, really rests on observations, whether consciously or not; and the only reason that this is not universally recognized is that it rests upon kinds of phenomena with which every man's experience is so saturated that he usually pays no particular attention to them. The data of metaphysics are not less open to

[2] See, for a contrast, Carnap's unfortunate definition of philosophy as 'the logic of science' in his 1937, § 72.

> observation, but immeasurably more so than the data, say, of the very highly observational science of astronomy… (1931, 6.2)[3]

Although overall Peirce's views were as close to Wittgenstein's as those of both were distant from the logical positivists and their theories, there is an important difference between both philosophers concerning the analysis of meaning. Peirce was generally interested in the connection between our concepts and praxis, including their practical effects, as a key to conceptual clarification and a better understanding of their meaning. But by proceeding in this way he risked extending the concept of meaning too far; he took a path that can easily lead us to confuse the cognitive and practical *effects* of meaning with meaning itself. For as we already saw, the cognitive meaning of a declarative sentence, seen as a combination of semantic-cognitive rules, works as a condition for the production of inferential awareness, which consists in the kind of *systemic openness* (allowing the 'propagation of content') that can produce an indeterminate number of subsequent mental states and actions.[4] Meaning as a verifiability rule is one thing; awareness of meaning and inferences that may result from this awareness, together with the practical effects of such inferences, may be a very different thing. Though they can be partially related, they should be distinguished. Hence, within our narrow form of inferentialism, we first have the inferences that construct meanings (like those of the identification rules of singular terms, the ascription rules of predicates, and the verifiability rules of sentences); then we have something usually beyond cognitive meaning, namely, the multiple inferences that enable us to gain something from our knowledge of meaning, along with the multiplicity of behavioral and practical effects that may result from them. Without this separation, we may even have a method that helps us clarify our ideas, but we will lack a boundary that can prevent us from extending the meanings of our expressions beyond a reasonable limit. For instance, the fact that something cannot be scratched helps to verify the assertion 'This is a diamond' (the hardest material), whereas the use of diamonds as abrasives will certainly be of little if any relevance for the explanation of the assertion's meaning. This is why I think that Wittgenstein, restricting cognitive meaning to a method of

---

[3] C. S. Peirce's view of metaphysics agrees with what is today the most accepted one (*Cf.* Loux 2001, ix). On Peirce's verificationism see also Misak 1995, Ch. 3. As I do, and following Peirce, Cheryl Misak favors a liberalized form of verificationism, opposed to the narrow forms advocated by the Vienna Circle.

[4] See my analysis of the form of semantic-cognitive rules in Chapter III, sec. 12, and considerations regarding the nature of consciousness in Chapter II, sec. 11.

verification, that is, to combinations of semantic rules able to make a proposition true, proposed a more adequate view of cognitive meaning and its truth.

Looking for a better example, consider the statement: (i) 'In October 1942 Chil Rajchman was arrested, put on a train, and deported to Treblinka.' This promptly leads us to the inference: (ii) 'Chil Rajchman died in a death camp.' However, his probable fate would not be part of the verifiability procedure of (i), but rather of the verification of statement (ii). Thus, although (ii) is easily considered a consequence of (i), its thought-content *isn't* a real constituent of the cognitive meaning, the thought-content-rule expressed by (i). Statement (ii) has its own verifiability procedures, even if its meaning is strongly associated with that of statement (i) since it is our main reason for being interested in this last statement. So, we could say that considering a statement *S*, there is something like a cloud of meanings surrounding its cognitive meaning, this cloud being formed by inferentially associated cognitive meanings of other statements with their own verifiability rules. But it is clear that this cloud of meaning does not properly belong to the cognitive meaning of *S* and should not be confused with it. In short: only by restricting ourselves to the constitutive verifiability procedures of a chosen statement are we able to restrict ourselves to the proper limits of its cognitive meaning.

Opposition to a reductionist replacement of metaphysics by science was also one reason why Wittgenstein didn't bother to make his principle formally precise, unlike positivist philosophers from A. J. Ayer to Rudolph Carnap. In saying this, I am not rejecting formalist approaches. I am only warning that such undertakings, if not well supported by a sufficiently careful pragmatic consideration of how language really works, tend to put the logical cart before the semantic horse. In this chapter, I want to show how unwise neglect of some very natural conceptual intuitions has frustrated most attempts by positivist philosophers to defend their own principle.

Having considered these differences, I want to start by examining some of Wittgenstein's remarks regarding the verifiability principle, in order to find a sufficiently adequate and reasonably justified formulation. Afterward, I will answer the main objections against the principle, demonstrating that they are much weaker than they seem at first glance.

2. Wittgensteinian semantic verificationism

Here are some of Wittgenstein's statements presenting the verifiability principle:

> Each sentence (*Satz*) is a signpost for its verification. (1984e: 150)
>
> A sentence (*Satz*) without any way of verification has no sense (*Sinn*). (1984f: 245)
>
> If two sentences are true or false under the same conditions, they have the same sense (even if they look different). (1984f: 244)
>
> To understand the sense of a sentence is to know how the issue of its truth or falsity is to be decided. (1984e: 43)
>
> Determine under what conditions a sentence can be true or false, then determine thereby the sense of the sentence. (This is the foundation of our truth-functions.) (1984f: 47)
>
> To know the meaning of a sentence, we need to find a well-defined procedure to see if the sentence is true. (1984f: 244)
>
> The method of verification is not a means, a vehicle, but the sense itself. Determine under what conditions a sentence must be true or false, thus determine the meaning of the sentence. (1984f: 226-7)
>
> The meaning of a sentence is its method of verification. (1980: 29)[5]

What calls attention to statements like these is their strongly intuitive appeal: they *seem* to be true. They satisfy our need for a methodological starting point that accords with our *common knowledge* beliefs. To a great extent, they even seem to corroborate Wittgenstein's controversial view, according to which philosophical theses should be ultimately trivial because they do no more than make explicit what we already know. They are what he would call 'grammatical sentences' expressing the rules grounding the linguistic practices that constitute our factual language. In the end the appeal to meaning verificationism involves what we might call a 'transcendental argument': we cannot conceive a different way to analyze the cognitive meaning of a declarative sentence, except by appealing to verifiability; hence, if we assume that cognitive meaning is analyzable, some form of semantic verificationism must be right.

---

[5] I believe that the germ of the verifiability principle is already present in aphorism 3.11 of the *Tractatus Logico-Philosophicus* under the title 'method of projection.' There he wrote: 'We use the perceptible sign of a sentence (spoken or written) as a projection of a possible state of affairs. The method of projection is the thinking of the sentence's sense.'

There are some points we can add. The first is terminological and was already extensively discussed in this book: we should not forget that the verifiability rule must be identified with the *cognitive content* of a declarative sentence. This cognitive content is what we could call, remembering our reconstruction of Frege's semantics, the *e-thought-content-rule* expressed by the declarative sentence (being also called the descriptive, informative or factual content of the sentence, if not its proposition or propositional content). A complementary point, already noted, is that we should never confuse cognitive content with grammatical meaning. If you do not know who Tito and Baby are, you cannot understand the cognitive meaning of the sentence 'Tito loves Baby,' even if you are already able to understand its grammatical meaning.

Another point to be emphasized is that the verifiability rule correctly understood as e-thought-content or proposition must include both, the verification *and* the falsification of the statement, since this rule can in itself be either true or false.[6] Wittgenstein was explicit about that: 'The method of verification is not a means, a vehicle, but the sense itself' (1984f: 226-7). The reason is easy to see: the verifiability e-thought rule either applies to the verifier as such – the truth-maker, which in the last chapter we usually and unequivocally identified with some cognitively independent *fact in the world* – which verifies the rule – or it *does not* apply to any expected verifier or fact in the world – which falsifies the rule. Consider, for example, the statement 'Frege was bearded.' Here the verifiability e-thought rule applies to a circumstantial fact that the rule is intended to apply to in a world that makes the rule effectively applicable, which means that the verifiability e-thought rule expressed by the statement is true. Consider, by contrast, the statement 'Wittgenstein was bearded': here the verifiability e-thought rule does not apply to the intended contextual fact in the world, since this fact does not exist, and that falsifies the statement. But then it is because the verifiability rule expressed by this statement is false, since it is inapplicable.

A final point concerns the reading of Wittgenstein's distinction between the verification of a *sentence* (*Satz*) and of a *hypothesis* (*Hypothese*), which he made in the obscure last chapter of his *Philosophical Remarks*. As he wrote:

> A hypothesis is a law for the building of sentences. One could say: a hypothesis is a law for the building of expectations. A sentence is, so to speak, a cut in our hypothesis in a certain place. (1984e XXII, sec. 228)

[6] This is why there is no falsifiability rule, as some authors like Michael Dummett have suggested (1993: 93).

In my understanding, the hypothesis is distinguished here mainly by being more distant from sensory-perceptual experience than what he calls a sentence. As a consequence, only the verification of a sentence (statement) is able to give us certainty. However, this does not mean that the verification of this sentence is infallible. Hence, when Wittgenstein writes that we can verify the truth of the sentence 'Here is a chair' by looking only at one *side* of the chair (1984e, Ch. XXII sec. 225), it is clear that we can increase our degree of certainty by adding new facets, aspects, modes of presentation, sub-facts. We could, e.g., look at the chair from other angles, or make tests to show what the chair consists of, whether it is solid enough to support a person, etc.

Thus, my take is that what he calls the *certainty* of a sentence is *only postulated as such* after we consider it sufficiently verified in the context of some linguistic practice. This is why things can be seen as certain and yet remain fallible, as practical certainties. By contrast, the verification of hypotheses, like sentences stating scientific laws, as this is realized only derivatively, gives us comparatively lower degrees of probability, though they can also be assumed as true.

3. Verifiability rule as a criterial rule

A more important point emphasized by Wittgenstein and ignored by others is that we usually have a choice of ways to verify a statement, each way constituting some different, more or less central aspect of its meaning. As he noted:

> Consideration of how the meaning of a sentence is explained makes clear the connection between meaning and verification. Reading that Cambridge won the boat race, which confirms that 'Cambridge won,' is obviously not the meaning, but is connected with it. 'Cambridge won' isn't the disjunction 'I saw the race or I read the result or...' It's more complicated. But if we exclude any of the means to check the sentence, we change its meaning. It would be a violation of grammatical rules if we disregarded something that always accompanied a meaning. And if you dropped all the means of verification, it would destroy the meaning. Of course, not every kind of check is actually used to verify 'Cambridge won,' nor does any verification give the meaning. The different checks of winning the boat race have different places in the grammar of 'winning the boat race.' (2001: 29)

Moreover:

> All that is necessary for our sentences to have meaning is that in some sense our experience *would* agree with them or not. That is: the immediate experience should verify only something of them, a *facet*. This picture is taken immediately from reality because we say 'This is a chair' when we see only a side of it. (1984f: 282, my italics)

In other words: one can verify through the direct observation of facts, that is, by seeing a Cambridge racing boat winning a race or by hearing the judge's confirmation, or both. These forms of verification are central to the meaning of 'Cambridge won the boat race.' It is worth remembering that even this direct observation of the fact is aspectual: each person at the boat race saw the fact from a different perspective, i.e., they saw and heard different sub-facts: different aspects (facets) of the same event. However, we also say that they all did *see* the grounding fact in the sense that they inferred its totality in the most direct way possible; this is why we can say that the fact-event of Cambridge winning**,** as a grounding fact, was also *directly* (that is, in the most direct possible way) experienced. In the same way, we are allowed to say that we see a ship on the sea (the inferred grounding fact), while what we phenomenally see is only one side of a ship (a given aspectual sub-fact).

However, often enough the way we can know the truth-value of a thought-content like that expressed by the sentence 'Cambridge won the boat race' is more indirect: someone can tell us, we can read this in the internet or in a magazine or we can see a trophy in the clubhouse… These ways are secondary, and for Wittgenstein they participate only secondarily in the sentence's meaning. Finally, they are causally dependent on more direct ways of knowing the truth-value, which are primary verifying criteria. If these first ways of verification did not exist, these dependent forms, being secondary criteria or mere symptoms, would lose their reliability and validity.

We can say that the verifiability rule applies when we achieve awareness of a fact, which means that we are in a position that allows us to make the relevant inferences from our factual knowledge. This awareness is the most direct when the criterial configuration (a configuration of p-properties or tropes) that satisfies the verifiability rule is at last partially constitutive of the grounding fact, for instance, when we observe a competition being won. But more often verification is indirect, namely, by means of secondary criteria or symptoms, often making the verifiability e-thought-content rule probably or even very probably true.

Criteria tend to be displayed in the form of criterial configurations, and such conditions can vary indeterminately. Thus, the verifiability rule is said to apply when a criterial configuration demanded by the semantic-

cognitive criterial rule is objectively given as belonging to objective facts as their constitutive tropical combinations and arrangements. Furthermore, concerning a basal e-thought-content, also a criterial rule seems to have as a minimum condition for satisfaction of some kind of *structural isomorphism* between, on the one hand, the interrelated internal elements originating as constituents of the thought-content-criterial-rule, and, on the other hand, the interrelated objective *elements* (objective tropical combinations) that make up the grounding fact in the world. This is what would constitute the isomorphism with the grounding fact. Since experience is always aspectual and often indirect, this also means that the dependent criterial configurations belonging to the rule must also show a structural isomorphism with aspectual configurations of independent or external criterial arrangements of tropes (given in the world and experienced by the epistemic subject). This generates what we could call isomorphic relations with a sub-fact (say, a ship on the sea seen from one side), and enables us to infer the whole grounding fact (say, a whole ship on the sea). I expect to say more about this complicated issue in the last chapter.[7]

As this reconstruction of Wittgenstein's views shows, a sentence's meaning should be constituted by a verifiability rule that usually ramifies itself, requiring the actual or possible fulfillment of a multiplicity of criterial configurations, allowing us to infer facts in more or less direct ways. Hence, there are definitional criterial configurations (primary criteria) such as, in Wittgenstein's example, those based on direct observation by a spectator at a boat race. But there are also an indefinite number of secondary criterial configurations depending on the first ones. They are secondary criteria or even symptoms, allowing us to infer that Cambridge (more or less probably) won the boat race, etc. Here too, we can say that the primary criteria have a definitional character: once we accept them as really given and we can agree on this, our verifiability rule should apply with practical certainty by defining the arrangement of tropes (fact) accepted as given. Moreover, we can treat secondary criteria (like reading about an event in a magazine) as less certain, though still very probable, while symptoms (like having heard about the event) make the application of a verifiability rule only more or less probable. Thus, if an unreliable witness tells us that Cambridge won, we can conclude that it is probable that Cambridge won. However, what makes this probability acceptable is, as we noted, that we are assuming it is backed by some

[7] A justified explanation of the resource to structural isomorphism will be given only in Chapter VI, sec. 2-5.

observation of the fact by competent judges and eye-witnesses, that is, by primary criterial configurations.

Investigating the structure of verifiability rules has some consequences for the much-discussed traditional concept of *truth-conditions*. The truth-condition of a statement S can be defined as the condition sufficient for its e-thought-content-rule actually be the case. The truth condition for the statement 'Frege had a beard' is the condition that he actually did have a beard. This means that the truth-condition of *S* is the *condition that a certain fact can be given as S's truth-maker, that is, as satisfying the verifiability rule for S*. The given truth-maker, the fact, is an *objective actualization* of the truth-condition. Thus, the so-called 'realist' view (in Michael Dummett's sense) is mistaken, since according to it a truth-condition of a statement could possibly be given without at least some conception of criterial configurations (tropical configurations that would possibly warrant its existence), and its related verifiability e-thought rules could to some extent be at least conceivable.

Now, considering our analysis of the identification rules of proper names (Appendix of Chapter I) and of the ascription rules of predicative expressions (Ch. II, sec. 6), we can consider the verifiability rule of a singular predicative statement to be a combination of both in a more explicit way. We can get an idea of this by examining a very simple predicative statement: 'Aristotle was bearded.' For this we have first as the definitional identification rule for Aristotle the same rule already presented at the beginning[8]:

> IR-Aristotle: The name 'Aristotle' is applicable *iff* its bearer is the human being who sufficiently and more than any other person satisfies the condition(s) of having been born in Stagira in 384 BC, son of Phillip's court physician, lived the main part of his life in Athens and died in Chalcis in 322 BC and/or was the philosopher who developed the main ideas of the Aristotelian opus. (Auxiliary descriptions may be helpful, though they do not belong properly to the definition…)

And for the predicative expression '…was bearded' we may formulate the following definitional ascription rule:

> AR-bearded: The predicate '…is bearded' is ascribable *iff* its bearer is a human being who has the tropes (properties) of facial hair growth on the chin and/or cheeks and/or neck.

[8] Appendix of Chapter I, sec. 1.

Now, as we already know, we first apply the identification rule of the singular term in order to identify the object, subsequently applying the ascription rule of the general term by means of which we select the tropical cluster of the object identified by the first rule. Not only are there many possible ways in which the identification rule and the ascription rule can be satisfied, there are *still more ways* of verification for the whole e-thought-content expressed by 'Aristotle was bearded.' One of them is by examining the well-known marble bust of Aristotle preserved in Athens, another is by accepting the recorded testimony of his contemporaries that has come down to us, and still another is by learning that most ancient Greeks (particularly among the peripatetics) customarily wore beards as a badge of manhood. All this makes possible the satisfaction of AR-bearded for that human being (the criterial configurations on the chin and cheeks are satisfied), in addition to the satisfaction of IR-Aristotle. As we noted, we assume this criterially-based verification as practically certain, which allows us to say we know that Aristotle was bearded, even if we are aware that this is only indirectly established as highly probable. We can summarize the applicability (or judgment or truth-attribution) of the basal e-thought-content verifiability rule to the grounding fact that Aristotle was bearded by means of the following schema:

> ├ [[IR-Aristotle is applicable to its bearer]AR-bearded is applicable to this same bearer].

These brief comments on verificationism *à la* Wittgenstein suggest the need for more intensive pragmatic research on ways of verification. As we noted, the structure of a verifiability rule is normally ramified, and its details should vary in accordance with the kind of statement that expresses it. A detailed pragmatic investigation of diversified forms of verifiability rules seems to me an important task that as far as I know has not been attempted until now. In what follows, I will not try to correct this limitation. I will restrict myself to answering the main objections to the verifiability principle, showing that they are products of misunderstanding.

4. Objection 1: The principle is self-refuting

The first and most notorious objection to the principle of verifiability is that it is self-defeating. The argument runs as follows. The principle of verifiability must be either *analytic* or *synthetic*.[9] If it is analytic it must be

[9] For my account of analyticity, see sec. 12 of the present chapter.

tautological, that is, non-informative. However, it seems clearly informative in its task of elucidating cognitive meaning. Furthermore, analytic statements are self-evident, and denying them is contradictory or inconsistent, which does not seem to be the case with the principle of verifiability. Therefore, the principle is synthetic. But if it is synthetic, it needs to be verifiable in order to have meaning. Yet, when we try to apply the principle of verifiability to itself we find that it is unverifiable. Hence, the principle is metaphysical, which implies that it is devoid of meaning. The principle is meaningless by its own standards; and one cannot evaluate meaningfulness based on something that is itself meaningless.

Logical positivists tried to circumvent that objection by responding that the principle of verifiability has no truth-value, for it is nothing more than a *proposal,* a *recommendation*, or a *methodological requirement*.[10] A. J. Ayer advocated this view by challenging his readers to suggest a more persuasive option (1992: 148). However, a reader with the opposite convictions could respond that he simply doesn't feel the need to accept or opt for anything of the kind... Moreover, the thesis that the principle is only a proposal appears to be clearly *ad hoc*. It goes against Wittgenstein's assumption that all we are doing is exposing the already given intuitions underlying our natural language, the general principles embedded in it. Consequently, to impose on our language a methodological rule that does not belong to it would be arbitrary and misleading as a means of clarifying meaning.[11]

My suggestion is simply to keep Wittgenstein's original insight, according to which a principle of verifiability is nothing but *a very general grammatical sentence stating the way all our factual language must work to have cognitive content to which a truth-value can be assigned*. Once we understand that the principle should make our pre-existing linguistic dispositions explicit, we are entitled to think that it must be seen as an analytic-conceptual principle. More precisely, this principle would consist in the affirmation of a hidden synonymy between the phrases 'meaning as the cognitive content (e-thought-content-rule or proposition) expressed by a declarative sentence' and 'the procedures (combinations of rules) by which we may establish the truth-value of this same cognitive content.' Thus, taking *X* to be any declarative sentence, we can define the *cognitive*

[10] This position was supported by A. J. Ayer, Rudolf Carnap, Herbert Feigl and Hans Reichenbach (*Cf.* Misak 1995: 79-80).

[11] Ayer's view wasn't shared by all positivists. Moritz Schlick, closer to Wittgenstein, defended the view according to which all that the principle of verifiability does is to make explicit the way meaning is assigned to statements, both in our ordinary language and in the languages of science (1936: 342 f.).

value of *X* by means of the following analytic-conceptual sentence stating the verifiability principle:

> VP (*Df.*): Cognitive meaning (e-thought-content) of a declarative sentence *X* = the verifiability rule for *X*.

Against this, a critic can react by saying that this claim to analytic identity lacks intuitive evidence. Moreover, if the principle of verifiability were analytic, it would be non-informative, its denial being contradictory or incoherent. However, it appears that VP says something to the effect that in principle it can be denied. It seems at least conceivable that the cognitive meaning of statement *X*, the thought-content expressed by it, isn't a verifiability rule.

My reaction to this objection is to recall that an analytic sentence does not need to be transparent; it does not need to be immediately seen as necessarily true, and its negation does not need to be clearly seen as contradictory or incoherent. Assuming that mathematics is analytic, consider the case of the following sentence: '3,250 + (3 . 896) = 11,876 ÷ 2.' At first glance, this identity neither seems to be necessarily true nor does its negation seem incoherent; but a detailed presentation of the calculation shows that this must be the case. We can regard it as a *hidden* analytic truth, at first view not graspable because of its derivative character and our inability to see its truth on the spot.

This can be suggested by means of a thought-experiment. We can imagine a person with a better grasp of arithmetic than ours. For a child, 2 + 3 = 5 can be analytically transparent, as it is for me. For me, 12 . 12 = 144 is also transparently analytic (or intuitively true), though not to a child who has just started to learn arithmetic. But 144 . 144 = 20,736 isn't transparently analytic for me, although it may be so for a person with greater arithmetical skill. Indeed, I would guess that some persons with great arithmetical skill (as in the case of some *savants*) can recognize at a glance the truth of the identity '3,250 + (3 . 896) = 11,876 ÷ 2.' This means that the boundary line between transparent and derived or non-transparent (but deductively achievable) analytic truths is movable, depending on our cognitive capacities and to some degree affected by training. Thus, from an epistemically neutral point of view, the two types are on the same level since for God (the only epistemic subject able to see all truths at a glance) analytic truths must *all* be transparent.

In searching for a better-supported answer, we can now distinguish between *transparent* and *non-transparent* analytic-conceptual knowledge.[12] The sentences 'A triangle has three sides,' 'Red is not green' and 'Three is greater than two' express transparent analytic knowledge, since these relations are self-evident and their negation clearly contradictory. But not all analytic sentences are so. Sentences about geometry such as the one stating the Pythagorean Theorem express (I assume) analytic truths in non-applied Euclidean geometry, although this isn't transparent for me. Non-transparent analytic knowledge is based on demonstrations whose premises are made up of transparent analytic knowledge, namely, analytic truths we can intuitively grasp.

The arithmetical and geometrical examples of analytic statements presented above are only *elucidative*, which can mislead us to think that they are informative in the proper sense of the word. This leads us to the suggestion that the principle of verifiability is nothing but a non-transparent, hidden analytic statement.

Against this last suggestion, one could still object that the principle of verifiability cannot be stated along the same lines as a mathematical or geometrical demonstration. After all, in the case of a proved theorem it is easy to retrace the path that leads to its demonstration; but there is no analogous way to demonstrate the principle of verifiability.

However, the key to an answer may be found if we compare the principle of verifiability with statements that at first glance do not seem to be either analytic or demonstrable. Close examination reveals that they are in fact only non-transparent analytic truths. A well-known statement of this kind is the following:

> The same surface cannot be simultaneously red all over and green all over (under the same conditions of observation).

This statement isn't analytically transparent. In fact, it has been regarded by logical positivists and even contemporary philosophers as a serious candidate for what might be called a synthetic a priori judgment (*Cf.* Bonjour 1998: 100 f.). Nevertheless, we can show that it is actually a hidden analytic statement. We begin to see this when we consider that it seems transparently analytic that (i) visible colors can occupy surfaces, (ii)

[12] This distinction is inspired by Locke's original distinction between intuitive and demonstrative knowledge. I do not use Locke's distinction because, as is well known, he questionably applied it to non-analytic knowledge. (*Cf.* Locke 1975, book IV, Ch. II, § 7)

different colors are things that cannot simultaneously occupy the same surface all over, and (iii) red and green are different colors. From this, it seems to follow that the statement (iv) 'The same surface cannot be both red and green all over' must be true. Now, since (i), (ii) and (iii) seem to be intuitively analytic, (iv) should be analytic too, even if not so intuitively clear.[13] Here's how this argument can be formulated in a standard form:

(1) Two different things cannot occupy the same place all over at the same time.
(2) A surface constitutes a place.
(3) (1, 2) Two different things cannot occupy the same surface all over at the same time.
(4) Colors are things that can occupy surfaces.
(5) (3, 4) Two different colors cannot occupy the same surface all over at the same time.
(6) Red and green are different colors.
(7) (5, 6) Red and green cannot occupy the same surface all over at the same time.

To most people, premises (1), (2), (4) and (6) can be understood (preserving the intended context) as definitely analytic. Therefore, conclusion (7) must also be analytic, even if it does not appear to be so.

The suggestion that I want to make is that the principle of verifiability is also a true, non-trivial and non-transparent analytic sentence, and its self-evident character may be demonstrated through an elucidation of its more transparent assumptions in a way similar to that of the above argument. Here is how it can be made plausible by the following 'cumulative' argument:

(1) Semantic-cognitive rules are criterial rules applicable to (or satisfied by) independent criteria that are tropical properties.
(2) Cognitive (descriptive, representational, factual…) meanings (e-thought-contents) of statements are constituted by proper combinations of (referential) semantic-cognitive rules applicable to real or only conceivable arrangements of tropical properties and their combinations called facts.

[13] Obviously, such an example can be decontextualized and therefore cheated in many ways. One could say: red and blue, for instance, can be blended to produce purple on the same surface, which is a bit like both colors… Like everything, examples can also be stolen and then used in the most inappropriate ways.

(3) The truth-determination of cognitive meanings or e-thought-content-rules of statements lies in the effective applicability of the proper combinations of semantic-cognitive criterial rules constitutive of them by means of their agreement (correspondence) with the arrangements and combinations of those tropical properties called real facts able to satisfy their criteria.
(4) Combinations of semantic-cognitive criterial rules expressible by statements are able to be true or false respectively by their effective applicability or non-applicability to their corresponding real or only conceivable facts, building in this way what we may call their e-thought-content *verifiability rules*.
(5) (1-4) The cognitive meanings of statements consist in their verifiability rules.

To my ears, at least, premises (1), (2), (3), and (4) sound clearly analytic, though conclusion (5) does not seem as clearly analytic. I admit that my view of these premises as analytic derives from the whole background of assumptions gradually reached in the earlier chapters of this book: it is analytically obvious to me that contents, meanings or senses are constituted by the application of rules and their combinations. It is also analytically obvious to me that the relevant rules are semantic-cognitive rules that can be applied in combination to form cognitive meanings or thought-contents expressible by declarative sentences. Moreover, once these combinations of rules are satisfied by the adequate criterial configurations formed by real facts understood as tropical arrangements, they allow us to see them as effectively applicable, that is, as having a verifying fact as their referent and truth-maker. Such semantic-criterial combinations of (normally implicit) cognitive rules, when judged as effectively applicable to their verifying facts, are called true, otherwise they are called false. And these semantic-criterial combinations of cognitive rules can also be called e-thoughts (e-thought-content-rules), propositional contents or simply *verifiability rules*.

I am aware that a few stubborn philosophers would still vehemently disagree with my reasoning, insisting that they have different intuitions originating from different starting points. After all I have said until now, I confess to be unable to help. To make things easier, I prefer to avoid discussion, invoking the words of an imaginary character from J. L. Borges: 'Their impurities forbid them to recognize the splendor of truth.'[14]

[14] From his magnificent short story, 'El Tintorero Enmascarado Hákim de Merv.'

5. Objection 2: A formalist illusion

Logic can be illuminating but also deceptive. An example is offered by A. J. Ayer's attempt to formulate a precise version of the principle of verifiability in the form of a criterion of *factual meaningfulness*. In his first attempt to develop this kind of verifiability principle, he suggested that:

> …it is the mark of a genuine factual proposition… that some experiential propositions can be deduced from it in conjunction with certain other premises without being deducible from these other premises alone. (1952: 38-39)

That is, it is conceivable that a proposition *S* is verifiable if together with the auxiliary premise *P1* it implies an observational result *O*, as follows:

1. *S*
2. *P1*
3. *O*

Unfortunately, it was soon noted that Ayer's criterion of verifiability was faulty. As Ayer himself recognized, his formulation was 'too liberal, allowing meaning to any statement whatsoever.' (1952: 11) Why? Suppose that we have as *S* the meaningless sentence 'The absolute is lazy.' Conjoining it with an auxiliary premise *P1*, 'If the absolute is lazy, then snow is white,' we can – considering that the observation that snow is white is true and that the truth of 'The absolute is lazy' cannot be derived from the auxiliary premise alone – verify the sentence 'The absolute is lazy.'

Now, the core problem with Ayer's suggestion (which was not solved by his later attempt to remedy it[15]) is this: In order to derive the observation that snow is white, he assumes that a declarative sentence (which he somewhat confusingly called a 'proposition') whose meaningfulness is questioned is already able to attain a truth-value. But meaningless statements cannot attain any truth-value: if a sentence has a truth-value, then it must also have a meaning, or, as I prefer to say, it must also express a propositional content as an e-thought verifiability rule that is true only as effectively applicable. By assuming in advance a truth-value for the sentence under evaluation, Ayer's principle implicitly begs the question, because if a statement must already have a sense in order to have

[15] The difficulty made him propose a more complicated solution that the logician Alonzo Church proved to be equally faulty (*Cf.* Church 1949).

a truth-value, it cannot be proven to be senseless. Moreover, he does not allow the empirical statement in question to reveal its proper method of verification or even if it has one.[16]

In fact, we cannot imagine any way to give a truth-value to the sentence 'The absolute is lazy,' even a false one, simply because it is a grammatically correct but cognitively meaningless word combination. As a consequence, the sentence 'If the absolute is lazy, then snow is white' cannot imply that the conclusion 'Snow is white' is true in conjunction with the sentence 'The absolute is lazy.' To make this obviously clear, suppose we replace 'The absolute is lazy' with the equally meaningless symbols @#$, producing the conjunction '@#$ & (@#$ → Snow is white).' We cannot apply a truth-table to show the result of this because @#$, just as much as 'the absolute is lazy,' expresses no proposition at all. Even if the statement 'Snow is white' is meaningful, we cannot say that this formula allows us to derive the truth of 'Snow is white' from 'The absolute is lazy,' because @#$, as a meaningless combination of symbols, cannot even be considered false in order to materially imply the truth of 'Snow is white.'

A. G. Hempel committed a similar mistake when he pointed out that a sentence of the form 'S ∨ N', in which S is meaningful, but not N, must be verifiable, in this way making the whole disjunction meaningful (1959: 112). Now, as we have seen, the real form of this statement is 'S ∨ @#$.' Obviously, we cannot apply any truth-table to this. In this case, only the verifiable S has meaning and allows verification, not the whole disjunction, because this whole cannot be called a disjunction. The true form of this statement, if we wish to preserve this title, is simply S.

I can develop the point further by giving a contrasting suggestion as a criterion of cognitive meaningfulness, more akin to Wittgenstein's views. Consider the sentence 'This piece of metal is magnetized.' The question of its cognitive meaningfulness suggests verifiability procedures. An affirmative answer results from the application of the following verification procedure that *naturally flows* from the statement 'This piece of metal is magnetized' conjoined with some additional information:

(1) This is a piece of metal (observational sentence).
(2) If a piece of metal is magnetized, it will attract other objects made of iron (a criterion for the ascription rule of '…is magnetized'),

[16] I am surely not the first to notice this flaw. See Barry Gower 2006: 200.

(3) This piece of metal has attracted iron coins, which remained stuck to it (observational application of the ascription rule's criterion to the object already criterially identified by the identification rule).
(4) (From 1 to 3): It is certainly *true* that this piece of metal is magnetized.
(5) If the application of the combination of semantic-cognitive rules demanded by a statement is able to make it true, then this combination must be its cognitive meaning (a formulation of the verifiability principle).
(6) (4 to 6): The statement '[It is certainly true that] this piece of metal is magnetized' is cognitively meaningful (it expresses an e-thought-content verifiability rule).

We can see that in cases like this the different possible verifying procedures *flow naturally from our understanding of the declarative sentence that we intend to verify, once the conditions for its verification are given.* However, in the case of meaningless sentences like 'The absolute is lazy' or 'The nothing nothings,' we can find no verification procedure following naturally from them, and this is the real sign of their lack of cognitive meaning. Ayer's statement 'If the absolute is lazy, then snow is white' does not follow naturally from the sentence 'The absolute is lazy.' In other words: the multiple ways of verifying a statement – themselves expressible by other statements – must contribute, in different measures, to make it fully meaningful; but they do this by *building* its cognitive meaning and not by being arbitrarily attached to the sentence, as Ayer's proposal suggests. They must be given to us intuitively as the declarative sentence's *proper* ways of verification. The neglect of real ways of verification naturally built into any genuine declarative sentence is the fatal flaw in Ayer's criterion.

6. Objection 3: Verificational holism

A sophisticated objection to semantic verificationism is found in W. V-O. Quine's generalization of Duhem's thesis, according to which it is impossible to confirm a scientific hypothesis in isolation, that is, apart from the assumptions constitutive of the theory to which it belongs. In Quine's concise sentence: '...our statements about the external world face

the tribunal of sense experience not individually but only as a corporate body.' (1951: 9)[17]

The result of this is Quine's *semantic holism*: our language forms a so interdependent network of meanings that it cannot be divided up into verifiability procedures explicative of the meaning of any isolated statement. The implication for semantic verificationism is clear: since what is verified must be our *whole* system of statements and not any statement alone, it makes no sense to think that each statement has an intrinsic verifiability rule that can be identified with a particular cognitive meaning. If two statements *S1* and *S2* can only be verified together with the system composed by {*S1*, *S2*, *S3*… *Sn*}, their verification must always be the same, and if the verifiability rule is the meaning, then all the statements should have the same meaning. This result is so absurd that it leaves room for skepticism, if not about meaning, as Quine would like, at least about his own argument.

In my view, if taken on a sufficiently abstract level, on which the concrete spatiotemporal confrontations with reality to be made by *each* statement are left out of consideration, the idea that the verification of any statement in some way depends on the verification of a whole system of statements – or, more plausibly, of a whole molecular sub-system – is very plausible. This is what I prefer to call *abstract or structural confirmational holism,* and this is what can be seriously meant in Quine's statement. However, his conclusion that the admission of structural holism destroys semantic verificationism, does not follow. It requires admitting that structural holism implies what I prefer to call a *performative, concrete or procedural verificational holism*, i.e., a holism regarding the concrete spatiotemporal verification procedures of individual statements, which are the only things really constitutive of their cognitive meanings. But this just never happens.

Putting things in a somewhat different way: Quine's holism has its seeds in the fact, well known by philosophers of science, that in order to be true the verification of an observational statement always depends on the truth of an undetermined multiplicity of assumed *auxiliary hypotheses* and *background knowledge*. Considered *in abstraction* from what we really do when we verify a statement, at least structural molecularism is true: verifications are interdependent. After all, our beliefs regarding any domain of knowledge are more or less interdependent, building a complex

[17] Later Quine corrected this thesis, advocating a verifiability molecularism restricted to sub-systems of language, since language has many relatively independent sub-systems. However, our counter-argument will apply to both cases.

network. But it is a wholly different matter if we claim that from formal or abstract confirmational holism, a performative procedural or verificational holism follows on a more concrete level. Quine's thesis is fallacious because, although at the end of the day a system of statements really needs to confront reality as a whole, in their concrete verification, its individual statements do not confront reality either conjunctively or simultaneously.

I can clarify what I mean with the help of a well-known example. We all know that by telescopic observation Galileo discovered the truth of the statement: (i) 'The planet Jupiter has four moons.' He verified this by observing and drawing, night after night, four luminous points near Jupiter, and concluding that these points were constantly changing their locations in a way that seemed to keep them close to the planet, crossing it, moving away and then approaching it again, repeating these same movements in a regular way. His conclusion was that these luminous points could be nothing other than moons orbiting the planet... Contemporaries, however, were suspicious of the results of his telescopic observation. How could two lenses magnify images without deforming them? Some even refused to look through the telescope, fearing it could be bewitched… Historians of science today have realized that Galileo's contemporaries were not as scientifically naive as they often seem to us.[18] As has been noted (Salmon 2002: 260), one reason for accepting the truth of the statement 'Jupiter has four moons' is the assumption that the telescope is a reliable instrument. But the reliability of telescopes was not sufficiently confirmed at that time. To improve the telescope as he did, Galileo certainly knew the law of telescopic magnification, whereby its power of magnification results from the focal length of the telescope divided by the focal length of the eyepiece. But in order to guarantee this auxiliary assumption, one would need to prove it using the laws of optics, still unknown when Galileo constructed his telescope. Consider, for instance, the fundamental law of refraction. This law was established by Snell in 1626, while Galileo's telescopic observations were made in 1610. With this addition, we can state in an abbreviated way the structural procedure of confirmation as it is known today and which I claim would be unwittingly confused by a Quinean philosopher with the concrete verification procedure. Here it is:

[18] I think Galileo's judges unwittingly did science a great favor by sentencing him to house arrest, leaving him with nothing to do other than concentrate his final intellectual energies on writing his scientific testament, the *Discorsi intorno a due nuove scienze*.

(I)
1. Repeated telescopic observation of four points of light orbiting Jupiter.
2. Law of magnification of telescopes.
3. Snell's law of refraction: $sin\theta_1/\sin\theta_2 = v1/v2 = \lambda1/\lambda2 = n2/n1$.
4. A telescope cannot be bewitched.
5. Jupiter is a planet.
6. The Earth is a planet.
7. The Earth is orbited by a moon.
8. (All other related assumptions.)
9. Conclusion: the planet Jupiter has at least four moons.

If Galileo did not have knowledge of premise 3, this only weakens the inductive argument, which was still strong enough to his lucid mind. From a Quinean verificationist holism, the conclusion, considering all the other constitutive assumptions, would be that the concluding statement 9 does not have a proper verification method, since it depends not only on observation 1, but also on the laws expressed in premises 2 and 3, the well-known premises from 4 to 7, and an undetermined number of other premises constitutive of our system of beliefs, all of them also having their verifiability procedures... As he wrote: 'our statements should face the tribunal of experience as a corporate body.' Indeed.

In this example, the problem with Quine's reasoning becomes clear. First, we need to remember that the premises belonging to confirmation procedures *are not simultaneously checked*. The conclusion expressed by statement 9 was actually verified only as a direct consequence of statement 1, resulting from the daily drawings made by Galileo based on his observations of variations in the positions of the four 'points of light' aligned near to Jupiter. However, Galileo did not simultaneously verify statement 2 when he made these observations, nor the remaining ones. In fact, as he inferred conclusion 9 from premise 1, he only *assumed* a previous verification of the other premises, as was the case with premise 2, which he verified as he learned how to build his telescope. Although he didn't have premise 3 as a presupposition, he had already verified or assumed as verified premises 2, 4, 5, 6, 7 and 8. Now, because in general the verifications of 2 to 8 are already made and presupposed during the verification of 9, it becomes clear that these verifications are totally independent of the actually performed verification of 9 by means of 1. The true form of Galileo's concrete verification procedure was much simpler than the abstract (holistic or molecularist) procedure of confirmation presented above. In a summarized form, it was:

1. Repeated telescopic observation of four points of light orbiting Jupiter.
2. Conclusion: the planet Jupiter has at least four moons.

Generalizing: If we call the statement to be verified *S*, and the statements of the observational and auxiliary hypotheses *O* and *A* respectively, the structure of the concrete verifiability procedure of *S* is *not*

$$\frac{O \qquad A_1 \,\&\, A_2 \ldots \,\&\, A_n}{S}$$

But simply:

$$\frac{O \qquad (\textit{Assuming}\ \text{the prior verification of}\ A_1 \,\&\, A_2 \ldots \,\&\, A_n)}{S}$$

This assumption of an *anterior* verification of auxiliary hypotheses in a way that might hierarchically presuppose sufficient background knowledge is what in practice makes all the difference, as it allows us to separate the verifiability procedure of *S* from the verifiability procedures of the involved auxiliary hypotheses and the many background beliefs which have been already successfully verified.

The conclusion is that we can clearly distinguish what verifies each auxiliary hypothesis. For example: the law of telescopic magnification was verified by very simple empirical measurements; and the law of refraction was established and verified later, based on empirical measurements of the relationship between variations in the angle of incidence of light and the density of the transmitting medium. Thus, while it is true that on an abstract level a statement's verification depends on the verification of other statements of a system, on the level of its proper cognitive and practical procedures, the successful verification of auxiliary and background statements is already assumed. This is what allows us to individuate the concrete verifiability procedure appropriate for a statement as what is actually being verified, *identifying it* with what we actually mean by the statement, thus with its proper cognitive meaning.

In the same way, we are able to distinguish the specific concrete modes of verification of each distinctive auxiliary or background statement, whose truth is assumed as verified before employing the verification procedure that leads us to accept *S* as true. This allows us to distinguish

and identify the concrete procedure or procedures whereby each statement of our system is cognitively verified, making the truth of abstract-structural holism irrelevant to the performative structure of semantic verificationism.

By considering all that is formally involved in confirmation, and by simultaneously disregarding the difference between what is presupposed and what is performed in the concrete spatiotemporal verification procedures, Quine's argument gives us the illusory impression that verification as such should be a holistic procedure. This seems to imply that the meaning of the statement cannot be identified with a verifiability procedure, since the meanings of the different statements are *multiple and diversified*, while the holistic confrontation of a system of beliefs with reality is *unique and as such undifferentiated.*

However, if we remember that each different statement must have a meaning of its own, it again becomes perfectly reasonable to identify the cognitive meaning of a statement with its verifiability rule! *For both the verifiability rule and the meaning are once more individuated together as belonging univocally to each statement,* and not to the system of statements or beliefs assumed in the verification. Molecular holism is true regarding the ultimate structure of confirmation. But it would be disastrous regarding meaning, since it would dissolve all meanings into one big, meaningless mush.

The inescapable conclusion is that Quine's verificational holism is false. It is false because the mere admission of formal holism, that is, of the fact that statements are in some measure inferentially intertwined with each other is insufficient to lead us to conclude that the verifiability rules belonging to these statements cannot be identified with their meanings because these rules cannot be isolated, as Quine suggested. Finally, one should not forget that in my example I gave only one way of verification for the statement 'The planet Jupiter has at least four moons.' Other ways of verification can be added, also constitutive of the meaning and enriching it and univocally related with the same statement.

Summarizing my argument: an examination of what happens when a particular statement is verified shows us that *even* assuming formal holism (which I think is generally correct, particularly in the form of a molecularism of linguistic practices), the rules of verifiability are distinguishable from each other *in the same measure* as the meanings of the corresponding statements – a conclusion that only reaffirms the expected correlation between the cognitive meaning of a statement and its method of verification.

7. Objection 4: Existential-universal asymmetry

The next well-known objection is that the principle of verifiability only applies conclusively to existential sentences, but not to universal ones. To verify an existential sentence such as 'At least one piece of copper expands when heated,' we need only observe a piece of copper that expands when heated. To conclusively verify a universal claim like 'All pieces of copper expand when heated' we would need to observe all the pieces of copper in the entire universe, including everything future and past, which is impossible. It is true that absolute universality is a fiction and that, when we talk about universal statements, we are always considering some limited domain of entities – some universe of discourse. But even in this case, the problem remains. In the case of metal expanding when heated, for instance, the domain of application remains much broader than anything we can effectively observe, making conclusive verification equally impossible.

A common reaction to this finding – mainly because scientific laws usually take the form of universal statements – is to ask whether it wouldn't be better to admit that the epistemic meaning of universal statements consists of falsifiability rules instead of verifiability rules… However, in this case existential sentences like 'There is at least one flying horse' would not be falsifiable, since we would need to search through an enormously vast domain of entities in the present, past and future in order to falsify it. Nonetheless, one could suggest that the meanings of universal statements were given by falsifiability rules, while the meanings of existential and singular statements would be given by verifiability rules. Wouldn't this be a more reasonable answer? (*Cf.* Hempel 1959)

Actually, though, I am inclined to think it would and could not do. We can, for example, falsify the statement 'All ravens are black' simply by finding a single white raven. In this case, we must simply verify the statement 'This raven is white.' In this way, the verifiability rule of this last statement is such that, if applied, it falsifies the statement 'All ravens are black.' But if the meaning of the universal statement may be a falsification rule, a rule able to falsify it, and the verifiability rule of the statement 'That raven is white' is the same rule that when applied falsifies the statement 'All ravens are black,' then – admitting that verifiability is the cognitive meaning of singular statements and falsifiability the meaning of the universal ones – it seems that we should agree that the statement 'All ravens are black' must be synonymous with 'That raven is white.' However, this would be absurd: the meaning of 'This raven is white' has almost nothing to do with the meaning of 'All ravens are black.'

The best argument I can think against falsifiability rules, however, is that they do not exist. As already noted, there seems to be *no proper falsifiability rule* for a statement, as there certainly is *no counter-assertoric force* (or a force proper to negative judgments, as once believed), *no rule of dis-identification* of a name, and *no rule for the dis-ascription or dis-application* of a predicate. This is because what satisfies a rule is a criterion and not its absence. – This is so even in those cases in which, by common agreement, the criterion is the absence of something normally expected, as in the case of a hole, e.g., if someone says: 'Your shirt has a hole in it,' or in the case of a shadow, in the statement 'This shadow is moving.' In such cases the ascription rule for '…has a hole' and the identification rule for 'This shadow' have what could be called 'negative criteria.' However, what needs to be satisfied or applied is the verifiability rule for the existence of a hole in the shirt, and not the falsifiability rule for the socially presentable shirt without a hole, since this would be the verifiability rule of a shirt that has no hole. And we use the verifiability rule for a moving shadow and not the falsifiability rule for the absence of a shadow. If I notice a curious moving shadow on a wall, I am verifying it; I am not falsifying the absence of moving shadows on the wall, even if the first observation implies the second.[19]

It seems, therefore, that we should admit that the cognitive meaning of a statement can only be its verifiability rule, applicable or not. But in this case, it seems at first view inevitable to return to the problem of the *inconclusive character* of the verification of universal propositions, leading us to the admission of a 'weak' together with a 'strong' form of verificationism as Ayer attempted to argue (1952: 37).

However, I doubt if this is the best approach to reach the right answer. My suggestion is that the inconclusiveness objection is simply faulty, since it emerges from a wrong understanding of the true logical form of universal statements; a brief examination shows that these statements are in fact *both probabilistic and conclusive*. Consider again the universal statement:

1. Copper expands when heated.

It is clear that its true logical form *is not*, as it seems:

[19] Michael Dummett viewed the falsification rule as the ability to recognize under what conditions a proposition is false (*Cf.* 1996: 62 f.). But this must be the same as the ability to recognize that the proposition isn't true, namely, that its verifiability rule isn't applicable, which presupposes that we know its criteria of applicability, being consequently able to recognize their absence.

2. [I affirm that] it is *absolutely certain* that all pieces of copper expand when heated,

whereby 'absolutely certain' means 'without possibility of error.' This logical pattern would be suitable for formal truths such as

3. [I affirm that] it is *absolutely certain* that 7 + 5 = 12,

because here there can be no error (except procedural error, which we are leaving out of consideration). However, this same form is not suitable for empirical truths, since we cannot be absolutely certain about their truth. The logical form of what we mean with statement (1) is a different one. This form is that of *practical certainty*, which can be expressed by

4. [I affirm that] it is *practically certain* that every piece of copper expands when heated,

where 'practically certain' means 'with a probability that is sufficiently high to make us disregard the possibility of error.' In fact, we couldn't rationally mean anything different from this. Now, if we accept this paraphrase, a statement such as 'Copper expands when heated' becomes *conclusively verifiable*, because we can clearly find inductive evidence protected by theoretical reasons that become so conclusive that we can be practically certain, namely, that we can assign the statement 'All pieces of copper expand when heated' a probability that is sufficiently high to make us very sure about it: we can affirm that we *know* its truth. In short: the logical form of an empirical universal statement – assuming there is some domain of application – is not that of a universal statement like '⊢ All S are P,' but usually:

5. [I affirm that] it is *practically certain* that all S are P.

Or (using a sign of assertion-judgment):

6. ⊢ It is *practically certain* that all S are P.

The objection of asymmetry has its origins in an internal transgression of the limits of language, in the case, the equivocal assimilation of the logical form of *empirical* universal statements in the logical form of *formal* universal statements (Chap. III, sec. 11). If the claims of empirical universal statement is nothing beyond a sufficiently high probability, this

is enough to make them conclusively verifiable. Hence, the cognitive meaning of an empirical universal statement can still be seen as its verifiability rule. Verification allows judgment; judgment must be treated as conclusive, and verification likewise.

8. Objection 5: Arbitrary indirectness

Another common objection is that the rule of verifiability of empirical statements requires taking as a starting point at least the *direct observation* of facts that are objects of a virtually interpersonal experience. However, many statements do not depend on direct observation to be true, as is the case with 'The mass of an electron is 9.109 x 10 kg raised to the thirty-first negative power.' Cases like this force us to admit that many verifiability rules cannot be based on more than *indirect observation* of the considered fact. As W. G. Lycan has noted, if we don't accept this, we will be left with a grotesque form of instrumentalism in which what is real must be reduced to what can be inter-subjectively observed and in which things like electrons and their masses do not exist anymore. But if we accept this, he thinks, admitting that many verifiability rules are indirect, how do we distinguish between direct and indirect observations? 'Is this not one of those desperately confusing distinctions?' (2000: 121 f.)

Here again, problems only emerge if we embark on the narrow formalist canoe of logical positivism, paddling straight ahead, only to tramp against the barriers of natural language with unsuitable requirements. Our assertoric sentences are inevitably uttered or thought in the contexts of language-games, practices, linguistic regions... The verification procedure must be adapted to the linguistic practice in which the statement is uttered. Consequently, the criterion to distinguish direct observation from indirect observation *should always be relative to the linguistic practice that we take as a model.* We can be misled by the fact that the most common linguistic practice is (A): *our wide linguistic practice of everyday direct observational verification*. The standard conditions for singling out this practice are:

> A possible interpersonal observation made by epistemic subjects under normal internal and external conditions and with unbiased senses of solid, opaque and medium-sized objects, which are close enough and under adequate lighting, all other things remaining the same.

This is how the presence of my laptop, my table and my chair are typically checked. Because it is the most usual form of observation, this practice is

seen as the archetypal candidate for the title of direct observation, to be contrasted with, say, indirect observation through perceptually accessible secondary criteria, as might be the case if we used mirrors, optical instruments, etc. However, it is an unfortunate mistake that some insist on using the widespread model (A) to evaluate what happens in other, sometimes very different, linguistic practices. Let us consider some of them.

I begin with (B): *the bacteriologist's linguistic practice*. Usually, the bacteriologist is concerned with the description of micro-organisms visible under his microscope. In his practice, when he sees a bacterium under a microscope, he says he has made a *direct* observation; this is his model for verification. But the bacteriologist can also say, for example, that he has verified the presence of a virus *indirectly*, due to changes he found in the form of the cells he saw under a microscope, even though for him viruses are not directly observable except under an electron microscope. If he does not possess one, he cannot make a direct observation of a virus. Almost nobody would say that the bacteriologist's procedures are all indirect unless they have in mind a comparison with our everyday linguistic practices (A). Anyway, although unusual, this would be possible. In any case, the right context and utterances clearly show what the speaker has in mind.

Let us consider now (C) *the linguistic practices of archaeology and paleontology*. The discovery of fossils is seen here as a direct way to verify the real existence of extinct creatures that died out millions of years ago, such as dinosaurs, since live observation is impossible, at least under any known conditions. But the archaeologist can also speak of indirect verification by comparison and contrast within his practice. So, consider the conclusion that hominids once lived in a certain place based only on damage caused by stone tools to fossil bones of animals that these early hominids once hunted and used for food or clothing. This finding may be regarded as resulting from an indirect verification in archaeological practice, in contrast to finding fossilized remains of early hominids, which would be considered a direct form of verification. Of course, here again, any of these verifications will be considered indirect when compared with verification by the most common linguistic observational practice of everyday life, that is (A). However, the context can easily show what sort of comparison we have in mind. A problem would arise only if the language used were vague enough to create doubts about the model of comparison employed.

If the practice is (D) one *of pointing to linguistically describable feelings*, the verification of a sentence will be called direct, albeit

subjective, if made by the speaker himself, while the determination of feelings by a third person, based on behavior or verbal testimony, will generally be taken as indirect (e.g., by non-behaviorists and many who accept my objections to the private-language argument). There isn't any easy way to compare practice (D) with the everyday practice (A) of observing medium-sized physical objects in order to say what is more direct, since they belong to two categorically different dimensions of verification.

My conclusion is that there is no real difficulty in distinguishing between direct and indirect verification, insofar as we have clarity about the linguistic practice in which the verification is being made, that is, about the model of comparison we have chosen (See Ch. III, sec. 7). Contrasted with philosophers, speakers normally share the contextually bounded linguistic assumptions needed for the applicability and truth-making of verifiability rules. To become capable of reaching agreement on whether a verificational observation or experience is direct or indirect, they merely need to be aware of the contextually established model of comparison that is being considered.

9. Objection 6: Empirical counterexamples

Another kind of objection concerns insidious statements that only seem to have meaning, but lack any effective verifiability rule. In my view, this kind of objection demands consideration on a case-by-case basis.

Consider, to begin with, the statement 'John was courageous,' spoken under circumstances in which John died without having had any opportunity to demonstrate courage, say, shortly after birth. (Dummett 1978: 148 f.) If we add the stipulation that the only way to verify that John was courageous would be by observing his behavior, the verification of this statement becomes practically (and very likely physically) impossible. Therefore, in accordance with the verifiability principle, this statement has no cognitive meaning, however, it still seems more than just grammatically meaningful.

The explanation is that under the described circumstances the statement 'John was courageous' only appears to have a meaning. It belongs to the sizable set of statements whose cognitive meaning is only apparent. Although the sentence has an obvious grammatical sense, given by the combination of a non-empty name with a predicate, we are left without any criterion for the application or non-application of the predicate. Thus, such a statement has no function in language, since it is unable to tell us anything. It is part of a set of statements such as 'The universe doubled in

size last night' and 'My brother died the day after tomorrow.' Although these statements may at first glance appear to have a sense, what they possess is no more than the expressive force of suggesting images or feelings in our minds. But in themselves, they are devoid of cognitive meaning since we cannot test or verify them.

Wittgenstein discussed an instructive case in his work *On Certainty*. Consider the statement 'You are in front of me right now,' said under normal circumstances for no reason by someone to a person standing before him. He notes that this statement only seems to make sense, given that we are able to imagine situations in which it would have some real linguistic function, for example, when a room is completely dark, so that it is hard for a person to identify another person in the room (1984a, sec. 10). According to him, we are inclined to imagine counterfactual situations in which the statement would or would not be true, and this invites us to project a truth-value into these possible situations and thus we will get the mistaken impression that the statement has some workable epistemic sense. Against this one could in a Gricean way still argue that even without any practical use the sentence has a literal assertoric sense, since it states something obviously true. However, this would be nothing but a further illusion: it *seems* to be obviously true only insofar as we are able to imagine situations in which it would make sense (e.g., exemplifying the evidential character of a perceptual assertion).

Finally, many statements are mediated and are only indirectly verifiable. Because of this, it is easy to make statements like 'The core of Jupiter is made of marschmallow,' and say that it is meaningful although unverifiable. However, we know that this statement is obviously false, and the method by which we falsify it is indirect since we cannot make a voyage to the center of Jupiter. We refute ramifications of the verification rule, which would deny our scientific conclusion that the core of this planet is made of water and helium and our awareness that marschmallow is made of milk and that there is no cow in Jupiter… These things show that the verifiability rule is inapplicable.[20]

What can we say of statements about the past or the future? Here too, it is necessary to examine them on a case-by-case basis. Suppose an expert says: 'Early Java man lived about 1 million years ago,' and this statement was fully verified by a reliable carbon dating applied to the fossilized skull. The direct verification of past events in the same way that we observe present events is practically (and it would seem physically)

[20] Another case is the verification of other minds. For an explanatory attempt, see my 2011, Ch. 4.

impossible. However, there is no reason to worry, since we are not dealing with the kind of verifiability rule adopted in standard practice (A). Here the linguistic practice assumed is (C), the archaeological, in which *direct* verification is made on the basis of verifiable empirical traces left by past events.

There are other, more indirect ways to verify past events. The sentence 'The planet Neptune existed before it was discovered' can be accepted as certainly true. Why? Because our knowledge of physical laws (which we trust as sufficiently verified), combined with information about the origins of our solar system, enables us to conclude that Neptune certainly existed a long time before it was discovered, and this inferential procedure is suitable as a form of verification. Finally, it is simply fallacious to say that since we can know about the past only by means of presently available evidence, we cannot say anything about the past, but only about our present, since the resource of present evidence can be the only natural and reliable way to speak about the past.

Very different is the case of statements about the past such as:

1. On that rock, an eagle landed exactly ten-thousand years ago.
2. Napoleon sneezed more than 30 times while he was invading Russia.
3. The number of human beings alive exactly 2,000 years ago was an odd number.

For such supposed thought-contents there are no empirical means of verification. Here we must turn to the old distinction between *practical*, *physical* and *logical* verifiability. Such verifications are not practically or technically achievable, and as far as I know, they are not even physically realizable (we will probably never be able to visit the past in a time-machine or travel through a worm-hole into the past in a spaceship). The possibility of verification of such statements seems to be only logical. But it is hard to believe that an *empirical* statement whose verifiability is only logical can be considered as having a non-logical cognitive sense (*Cf.* Reichenbach 1953: sec. 6).

To explain this point better: it seems that the well-known distinction between *logical*, *physical* and *practical* forms of verifiability exerts influence on meaningfulness depending on the respective fields of verifiability to which the related statements belong. Statements belonging to a formal field need only be formally verifiable to be fully meaningful: the tautology $(P \rightarrow Q) \leftrightarrow (\sim P \vee Q)$, for instance, is easily verified by the truth-table applying the corresponding logical operators. But statements

belonging to the empirical domain (physical and practical) must be not only logically, but also at least in principle empirically verifiable in order to have real cognitive meaning. As a consequence, an empirical statement that is only logically verifiable must be devoid of cognitive significance. This seems to be the case with a statement such as 'There is a nebula that is moving away from the earth at a speed greater than the speed of light.' Although logically conceivable, this statement is empirically devoid of sense, insofar as it is impossible according to relativity theory. Similarly, in examples (1), (2) and (3), what we have are empirical statements whose verification is empirically inconceivable. Consequently, although having grammatical and logical meaning and eliciting images in our minds, these statements *lack any distinctive cognitive value, for we don't know what to make of them*. Such statements aren't able to perform the specific function of an empirical statement, which is to be *able* to truly represent an actual state of affairs. We do not even know how to begin the construction of their proper verifiability rules. All that we can do is to imagine or conceive the situations described by them, but we know of no rule or procedure to link the conceived situation to something that possibly exists in the real world. Although endowed with grammatical and some expressive meaning, they are devoid of genuine cognitive meaning. Finally, we must remember that we are free to reformulate statements (1), (2) and (3) as meaningful empirical possibilities. For instance: (2') 'Maybe (it is possible that) Napoleon sneezed more than 30 times when he was invading Russia.' Although not very dissimilar to (2), this modal statement is verifiable as true by means of its coherence with our belief-system.

Also unproblematic is the verificational analysis of statements about the future. The great difference here is that in many cases direct verification is practically possible. Consider the sentence (i) 'It will rain in Caicó seven days from now.' When a person seriously says something of this sort, what he usually means is (ii) *'Probably* it will rain in Caicó seven days from now.' And this probability sentence can be conclusively verifiable, albeit indirectly, by a weather forecast. Thus, we have a verifiability rule, a cognitive meaning, and the application of this rule gives the statement a real degree of probability. However, one could not in anticipation affirm (iii) 'It *certainly* will rain within seven days.' Although there is a direct verifiability rule – watch the sky for seven days to determine if the thought-content is true or false – it has the disadvantage that we will only be able to apply it if we wait for a period of time, and we will only be able to affirm its truth (or deny it) within the maximal period of seven days. It is true that we could also use this sentence in certain situations, for example, when making a bet about the future. But in this

case, we would not affirm (iii) from the start since we cannot apply the rule in anticipation. In this case, what we mean with sentence (i) can in fact only be (iv) *'I bet that* it will rain in Caicó seven days from now.' Lacking any empirical justification, the bet has again only an expressive-emotive meaning and no truth-value.

A similar statement is (v) 'The first baby to be born on Madeira Island in 2050 will be female,' which has a verifiability rule that can only be applied at a future point in time. This sentence lacks a practical meaning insofar as we are unable to verify and affirm it at the present moment; right now this sentence, though expressing a thought-content – since it has a verifiability rule whose application can be tested in the future – is able to have a truth-value, but cannot receive it until later. Nonetheless, in a proper context this sentence may also have the sense of a guess: (vi) *'I guess* that the first baby to be born…' or (vii) a statement of possibility regarding the future *'It is possible that* the first baby to be born…' In these cases, we are admitting that the sentence has a cognitive meaning since all we are saying is that it has an observational verifiability rule that can be applied (or not), although only in the future. Sentence (v) will only be meaningless if understood as an *affirmation* of something that is not now the case but will be the case in the year 2050, for in order to be judged to be true this affirmation requires awareness of the effective applicability of the verifiability rule generally based on its real application. (*Cf.* Ch. IV, sec. 36) When we consider what is really meant by statements regarding future occurrences, we see that even in these cases verifiability and meaning go together.

Now consider the statement (viii): 'In about eleven billion years the Sun will expand and engulf Mercury.' This statement in fact only means *'Very probably* in about eleven billion years the Sun will expand and engulf Mercury,' This probabilistic prediction can be inferentially verified today, based on what we know of the fate of other stars in the universe that resemble our Sun but are much older, and this inferential verification constitutes its cognitive meaning.

Jeopardizing positivist hopes, I conclude that there is no general formula specifying the general form of verifiability procedures. Statements about the future can be physically and to some extent also practically verifiable. They cannot make sense as warranted assertions about actual states of affairs since such affirmations require the possibility of present verification. Most of them are concealed probability statements. The kind of verifiability rule required depends on the utterance and its insertion in the linguistic practice in which it is made, only then showing clearly what

it really means. Such things are what may lead us to the mistaken conclusion that there are unverifiable statements with cognitive meaning.

Finally, a word about ethical statements. Positivist philosophers have maintained that they are unverifiable, which has led some to adopt implausible emotivist moral theories. Once again, we find the wrong attitude. I would rather suggest that ethical principles can be only more or less plausible, like metaphysical statements and indeed like any philosophical statement. They have the form: 'It is plausible that *p*,' and as such they are fully verifiable. They cannot be decisively affirmed because we are still unable to state them in adequate ways or make them sufficiently precise, since we lack consensual agreement regarding their most adequate formulation and verifiability rules.

10. Objection 7: Formal counterexamples

The verificationist thesis is naturally understood as extendable to the statements of formal sciences. In this case, the verifiability rules or procedures that demonstrate their formal truth constitute a form of cognitive content deductively, within the assumed formal system in which they are considered. A fundamental difference with respect to empirical verification is that in the case of formal verification, *to have a verifiability rule is the same thing as being definitely able to apply it, since the criteria ultimately to be satisfied are the own axioms already assumed as such by the chosen system.*

A much discussed counterexample is Goldbach's conjecture. This conjecture (G) is usually formulated as:

> G: Every even number greater than 2 can be expressed as the sum of two prime numbers.

The usual objection is that this mere conjecture has cognitive meaning. It expresses a thought-content even if we never manage to prove it, even if a procedure for formal verification of G has not yet been developed. Therefore, its significance cannot be equated with a verifiability procedure.

The answer to this objection is quite simple and stems from the perception that Goldbach's conjecture is what its name says: a mere *conjecture*. Well, what is a conjecture? It's not an affirmation, a proven theorem, but rather the recognition that an e-thought-content-rule has *enough plausibility* to be taken seriously as possibly true. One would not

make a conjecture if it seemed fundamentally improbable. Thus, the true form of Goldbach's conjecture is:

It is plausible that G.

But 'It is plausible that G,' that is, '[I state that] it is plausible that G,' or (using a sign of assertion) ' ⊢It is plausible that G,' is something other than

I state that G (or ⊢G),

which is what we would be allowed to say if we wanted to state Goldbach's *proved theorem*. If our aim were to support the statement 'I state that G,' namely, an affirmation of the truth of Goldbach's theorem as something cognitively meaningful, the required verifiability rule would be *the whole procedure for proving the theorem*, and this we simply do not have. In this sense, G is cognitively devoid of meaning. However, the verifiability rule for ascribing mere plausibility is far less demanding than the verifiability rule able to demonstrate or prove G, and we have indeed applied this rule many times.

The plausibility ascription is '[I state that] it is plausible that G,' whereby the verifiability rule consists in something much weaker, namely, a verification procedure able to *suggest* that G could be proved. This verification procedure does in fact exist. It consists simply in considering random examples, such as the numbers 4, 8, 12, 124, etc., and showing that they are always the sum of two prime numbers. This verifiability rule not only exists, up until now it has been confirmed without exception for every even natural number ever considered! This is the reason why we really do have enough support for Goldbach's conjecture: it has been fully verified *as a conjecture*. If an exception had been found, the conjecture would have been proved false, for this would be incompatible with the truth of '[I state that] it is plausible that G' and would from the start be a reason to deny the possibility of Goldbach's conjecture being a theorem.

Summarizing: in itself the conjecture is verifiable and – as a conjecture – *has been definitely verified*: It is simply true that G is highly plausible. And this justifies its cognitive meaningfulness. What remains beyond verification is the statement affirming the necessary truth of G. And indeed, this statement doesn't really make sense; it has no cognitive content since it consists in a proof, a mathematical procedure to verify it, which we do not have. The mistake consists in the confusion of the statement of a mere conjecture that is true with the 'statement' of a theorem that does not exist.

A contrasting case is Fermat's last theorem. Here is how this theorem (F) is usually formulated:

> F: There are *no* three positive integers x, y and z that satisfy the equation $x^n + y^n = z^n$, if n is greater than 2.

This theorem had been only partially demonstrated up until 1995 when Andrew Wiles finally succeeded in working out a full formal proof. Now, someone could object here that even before Wiles' demonstration, F was already called 'Fermat's theorem.' Hence, it is clear that a theorem can make sense even without being proved!

There are, however, two unfortunate confusions in this objection. The first is all too easy to spot. Of course, Fermat's last theorem has a grammatical sense: it is syntactically correct. But it would be an obvious mistake to confuse the grammatical meaning of F with its cognitive meaning as a theorem. Also an absurd identity, for instance, 'Napoleon is the number 7,' has a grammatical sense.

The second confusion concerns the fact that the phrase 'Fermat's theorem' isn't appropriate at all. We equivocally used to call F a 'theorem' because before his death Fermat wrote that he had proved it, but couldn't put this proof on paper since the margins of his notebook were too narrow…[21] For these reasons, we have here a misnamed opposite of 'Goldbach's theorem.' Although F was called a theorem, it was in fact only a conjecture of the form:

> [I state that] it is plausible that F.

It *was* a mere conjecture until Wiles demonstrated F, only then effectively making it a true theorem. Hence, before 1995 the cognitive content that could be given to F was actually '[I state that] it is plausible that F,' a conjecture that was initially demonstrated by the fact that no one had ever found numbers x, y and z that could satisfy the equation. Indeed, the cognitive meaning of the real theorem F, better expressed as 'I state that F' or '$\vdash$ F' (a meaning that very few really know in its entirety), should include the demonstration or verification found by Wiles, which is no more than the application of an exceptionally complicated combination of mathematical rules.

[21] Today we know that Fermat was only joking since the mathematics of his time did not provide the means to prove his conjecture.

Some would complain that if this is the case, then only very few people really know the cognitive meaning of Fermat's last theorem. I agree with this, though seeing no reason to complain. The cognitive content of this theorem, its full thought-content, like that of many scientific statements, is really known by very few people indeed. What most of us know is only the weak conjecture falsely called 'Fermat's last theorem'. We have applied F to some numbers without finding any exception.

Finally, there are phrases like (i) 'the less rapidly convergent series.' For Frege, this definite description has sense but not reference (1892: 28). We can add that there is a rule that allows us to always find series that are less convergent than any given one, making them potentially infinite. We can state this rule as L: 'For any given convergent series, we can always find a less rapidly convergent one.' Since L implies the truth of statement (ii) 'There is no less rapidly convergent series,' we conclude that (i) has no referent. Now, what is the identification rule of (i)? What is the sense, the meaning of (i)? One answer would be to say that it is given by failed attempts to create a less rapidly convergent series ignoring L. It would be like the meaning of any mathematical falsity. For instance, the identity (iii) 321 + 427 = 738 is false. Now, what is its meaning? A temptation is to classify it as senseless. But if it were senseless, it would not be false. Consequently, I suggest that its sense resides in the failed usual ways to verify it, which leads to the conclusion that this is a false identity. It seems reasonable to conclude that it is such an external operation that gives a kind of cognitive sense to a false identity. The same holds regarding false statements like 3 > 5. They express misrepresentations, incongruities demonstrating failed attempts to apply rules in the required ways.

11. Objection 8: Skepticism about rules

In his *Philosophical Investigations,* Wittgenstein formulated a skeptical paradox (1984c, sec. 201) that endangers the possibility of an ongoing *common interpretation* of rules and, consequently, the idea that our language may work as a system of rules responsible for meaning. Solving this riddle interests us here because if the argument is correct, it seems to imply that it is a mistake to accept that there are verifiability rules consisting in the cognitive meanings of sentences.

Wittgenstein's paradox results from the following example of rule-following. Let's say that a person learns a rule to add 2 to natural numbers. If you give him the number 6, he adds 2 and writes the number 8. If you give him the number 173, he adds 2, writing the number 175... But imagine that for the first time he is presented with a larger number, say the

number 1,000, and that he then writes the number 2,004. If you ask why he did this, he responds that he understood that he should add 2 up to the number 1,000, 4 up to 2,000, 6 up to 3,000, etc. (1984c, sec. 185).

According to Saul Kripke's dramatized version of the same paradox, a person learns the rule of addition, which works well for additions with numbers below 57. But when he performs additions with larger numbers, the result is always 5. So for him 59 + 67 = 5… Afterward, we discover that he understood 'plus' as the rule 'quus,' according to which 'x quus y = $x + y$ if $\{x, y\} < 57$, otherwise 5' (1982: 9). If questioned why he understood addition in this strange way, he answers that he found this the most natural way to understand the rule.

Now, what these two examples suggest is that a rule can always be interpreted differently from the way it was intended, no matter how many specifications we include in our instructions for using the rule, since these instructions can also be differently interpreted… As Kripke pointed out, there is *no fact of the matter* that forces us to interpret a rule in a certain way rather than in any other. The consequence is that we cannot be assured that everyone will follow our rules in an expected similar way, or that people will continue to coordinate their actions based on them. And as meaning depends upon following rules, we cannot be certain about the meanings of the expressions we use. How could we be certain, in the exemplified cases, of the respective meanings of 'add two' and 'plus'? However, if we accept that there can be no rules and therefore no meanings, then there could be no riddle since we would not be able to meaningfully formulate the riddle.

Wittgenstein and later Kripke attempted to find a solution to the riddle. Wittgenstein's answer can be interpreted as saying that we follow rules blindly, as a result of training (custom) regarding the conventions of our social practices and institutions belonging to our way of life (1984c sec. 198, 199, 201, 219, 241). Kripke's answer follows a similar logic: according to him, following a rule isn't justified by truth-conditions derived from their correct interpretation in a correspondential (realist) way, a solution that Wittgenstein tried in his *Tractatus*. Instead, Kripke thinks that for the later Wittgenstein correspondence is replaced by verification, so that instead of truth-conditions what we have are *assertability conditions* justified by practical interpersonal utility (1982: 71-74, 77, 108-110). These assertability conditions are grounded on the fact that any other user in the same language community can assert that the rule follower 'passes the tests for rule following applied to any member of the community' (1982: 110).

Notwithstanding, both answers are clearly wanting. They offer a description of *how* rules work, leaving unexplained *why they must work*. Admittedly, the simple fact that in our community we have so far openly coordinated our linguistic activity according to rules does not imply that this coordination *has to* work this way, nor does it imply that it should *continue* to work this way. Kripke's answer has in my view an additional burden. It overlooks the fact that assertability conditions must include the satisfaction of truth-conditional correspondential-verificational conditions, only adding to the explanation of the common interpretation of rules an interpersonal social layer.

For my part, I have always believed that the 'paradox' should have a more satisfactory solution. A central point can be seen as in some way already disclosed by Wittgenstein, namely, that we learn rules in a similar way because *we share a similar human nature modeled in our form of life*. It seems clear that this makes it easier for us to interpret the rules we are taught in the same manner, suggesting that we must also be naturally endowed with innate, internal corrective mechanisms able to reinforce consistent, conforming behavior. (Costa 1990: 64-66)

Following this path, we are led to the decisive solution of the riddle, which I think we owe to Craig DeLancey (2004). According to him, we are biologically predisposed to construct and interpret statements *in the most economical (or parsimonious) way possible*. Or, as I prefer to say, we are innately disposed to put in practice the following principle of simplicity:

> PS: We should formulate and interpret a semantic rule in the simplest way possible.

Because of this shared principle derived from our inborn nature as rule followers, we prefer to maintain the interpretation of the rule 'add 2' in its usual form, instead of complicating it with the further condition that we should add twice two after each thousand. And because of the same principle, we prefer to interpret the rule of addition as a 'plus' instead of a 'quus' addition, because with the 'quus' addition we would complicate the interpretation by adding the further condition that any sum with numbers above 57 would give as a result the number 5. Indeed, it is the application of this principle of simplicity that is the 'fact of the matter' not found by Kripke, which leads us to interpret a rule in one way instead of another. It allows us to harmonize our interpretations of semantic rules, thus solving the riddle. Furthermore, DeLancey clarifies 'simplicity' by remarking that non-deviant interpretations are formally *more compressible* than deviant

interpretations like those considered by Wittgenstein and Kripke. Moreover, a Turing machine would need to have a more complex and longer program in order to process these deviant interpretations...

One might ask: what warrants assuming the long-term consistency of human nature across the entire population or that we are innately equipped to develop such a heuristic principle of simplicity? The obvious answer lies in the appeal to Darwinian evolution. Over long periods of time, a process of natural selection has harmonized our learning capacities around the principle of simplicity and eliminated individuals with deviant, less practical dispositions. Thus, we have a plausible explanation of our capacity to share a sufficiently similar understanding and meaning of semantic rules. If we add to this the assumption that human nature and recurring patterns in the world will not change in the future, we can be confident in the expectation that people will not deviate from the semantic rules they have learned. Of course, underlying this last assumption is Hume's much more defiant criticism of induction, which might remain a hidden source of concern. But this is a further issue that goes beyond our present concerns (for a plausible approach see the Appendix of the present chapter).[22]

Summarizing: Our shared interpretation of learned rules only seems puzzling if we insist on ignoring the implications of the theory of evolution, which supports the principle of simplicity. By ignoring considerations like these, we tend to ask ourselves (as Wittgenstein and Kripke did) how it is possible that these rules are and continue to be interpreted and applied in a similar manner by other human beings, losing ourselves within a maze of philosophical perplexities. For a similar reason, modern pre-Darwinian philosophers like Leibniz wondered why our minds are such that we are able to understand each other, appealing to the Creator as producing the necessary harmony among human souls. The puzzle about understanding how to follow rules arises from this same old perplexity.

---

[22] Curiously, in his book Kripke considers the criterion of simplicity, but repudiates it almost fortuitously for the reason that 'although it allows us to choose between different hypotheses, it can never tell us what the competing *hypotheses* are' (1982: 38). However, what the competing hypotheses – call them the rules *x* and *y* – ultimately are, is a metaphysically idle question, only answerable by God's omniscience, assuming that the concept of omniscience makes any sense. The real paradox appears only when we can state it in the form of comparable hypotheses like 'plus' *versus* 'quus,' and it is to just such cases that we apply the principle of simplicity.

12. Quine's objections to analyticity

Since I am assuming that the verifiability principle is an analytic-conceptual statement, before finishing I wish to say a word in defense of analyticity. I am satisfied with the definition of an analytic proposition as *the thought-content expressed by a statement whose truth derives from the combination of its constitutive unities of sense*. This is certainly the most common and intuitively acceptable formulation. However, W. V-O. Quine would reject it because it seems to be based on an overly vague and obscure concept of meaning.

The usual answer to this criticism is that there is really nothing overly vague or obscure in the concept of meaning used in our *definiens*, except from Quine's own scientistic-reductionist perspective, which tends to confuse expected vagueness with lack of precision and obscurity (*See* Grice & Strawson 1956: 141-158; Swinburne 1975: 225-243). Philosophy works with concepts such as meaning, truth, knowledge, good… which are in some measure polysemic and vague, as much so as the concepts used in countless attempts to define them. In my judgment, the effort to explain away such concepts only by reason of their vagueness (or supposed obscurity) betrays an impatient positivist-scientistic mental disposition, which is anti-philosophical *par excellence* (which doesn't mean to indulge the opposite: a methodology of hyper-vagueness or unjustified obscurity).

Having let out of consideration the above definition, Quine tried to define an analytic sentence in a Fregean way, as a sentence that is either tautological (true because of its logical constants) or can be shown to be tautological by the replacement of its non-logical terms with cognitive synonyms. Thus, the statement (i) 'Bachelors are unmarried adult males' is analytic, because the word 'bachelor' is a synonym of the phrase 'unmarried adult male,' which allows us by the substitution of synonyms to show that (i) means the same thing as (ii): 'Unmarried adult males are unmarried,' which is a tautology. However, he finds the word 'synonym' in need of explanation. What is a synonym? Quine's first answer is that the synonym of an expression is another expression that can replace the first in all contexts *salva veritate*. However, this answer does not work in some cases. Consider the phrases 'creature with a heart' and 'creature with kidneys.' They are not synonymous, but are interchangeable *salva veritate*, since they have the same extension. In a further attempt to define analyticity, Quine makes an appeal to the modal notion of necessity: 'Bachelors are unmarried males' is analytic if and only if 'Necessarily, bachelors are unmarried males.' But he also sees that the usual notion of necessity does not cover all cases. Phrases like 'equilateral triangle' and

'equiangular triangle' necessarily have the same extension, but are not synonyms. Consequently, we must define 'necessary', in this case, as the specific necessity of analytic statements, in order for the concept to apply in all possible circumstances... However, the 'necessity of analyticity' is an obscure notion, if it really exists. Dissatisfied, Quine concludes that his argument to explain analyticity 'has the form, figuratively speaking, of a closed curve in space.' (Quine 1951: 8)

A problem emerges from Quine's implicit assumption that a word should be defined with the help of words that do not belong to its specific conceptual field. Thus, for him, the word 'analyticity' should not be defined by means of words like 'meaning,' 'synonymy,' 'necessity'… which just as much as 'analyticity' seem too near and unspecific in their meaning to be trusted in the construction of an adequate definition. Nonetheless, when we consider the point more carefully, we see that the words belonging to a *definiens* should be sufficiently close in their meanings to the *definiendum*, simply because in any real definition the terms of a *definiens* must belong to the same semantic field as its *definiendum*, notwithstanding the element of vagueness. This is why, in order to define a concept-word from ornithology, we would not use concepts from quantum mechanics, and vice-versa. These conceptual fields are too distant from each other. Because of this, we define 'arthropod' as an invertebrate animal having an exoskeleton, all these terms being biological, which does not compromise the definition. And considering the abstractness of the semantic field, a kindred level of vagueness can be expected. Hence, there is nothing especially wrong in defining analyticity using correspondingly vague words belonging to the same conceptual field, like 'meaning' and 'synonymy,' refraining from further elucidation.

A more specific and more serious objection is that Quine's attempt to define synonymy simply took a wrong turn. Since there is probably no proper necessity of analyticity, the lack of synonymy of expressions that necessarily have extensions like 'equilateral triangle' and 'equiangular triangle' remains unexplained.

My alternative proposal consists simply in beginning with the dictionary definition according to which:

> Two words or phrases are synonymous when they have the same or nearly the same meaning as another word or phrase in the same language.[23]

[23] Oxford Dictionaries.

Translating this into our terms, this means that *any expressions A and B are (cognitively) synonymous if their semantic-cognitive rules (their expressed concepts) are the same or almost the same*. This can be tested by *adequate definitions (analyses)* expressing the criteria for the application of those rules so that when these rules are really the same, the synonymous expressions will be called *precise synonyms*. However, precise synonyms are difficult to find. Consider, for instance, the words 'beard' and 'facial hair.' These words are called synonymous because they express a similar semantic-cognitive rule. A 'beard' is defined by dictionaries as 'a growth of hair on the chin and lower cheeks of a man's face' and this is considered sufficiently similar to the expression 'facial hair.' However, the two terms are not precisely synonymous, because a human being with hair on the forehead has facial hair but no beard. Diversely, the word 'chair' and the expression 'a non-vehicular seat provided with a backrest and made for use by only one person at a time' can be seen as precise synonymous, because the latter is simply the real definition of the former. The expressions 'creature with a heart' and 'creature with a kidney', on the other hand, are not synonymous, because they express different semantic-cognitive rules, the first defined as a creature with an organ used to pump blood, the second defined as a creature with an organ used to clean waste and impurities from blood. Even if approximate in meaning, the expressions 'equilateral triangle' and 'equiangular triangle' are surely not precisely synonymous for the reason already considered: the first is defined as a triangle whose three sides are equal, while the second is defined as a triangle whose three internal angles are congruent with each other and are each 60°. Hence, we can replace Quine's flawed definition of analyticity with the following more adequate definition using the concept of precise synonymy:

> A statement *S* is analytic (*Df*): It can generate a tautology by means of substitution of precise cognitive synonyms, namely, of real definitions expressing the same semantic-cognitive criterial rules.

The statement 'The cognitive meaning (e-thought-content) of a declarative sentence *X* = the verifiability rule for *X*' is analytic because the semantic-cognitive rules on each side of the identity sign are identical.

A complementary point supported by Quine is that, contrary to what is normally asserted, there is no definite distinction between empirical and formal knowledge. What we regard as analytic sentences can always be falsified by greater changes in our more comprehensive system of beliefs.

Even sentences of logic such as the excluded middle can be rejected, as occurs in some interpretations of quantum physics.

Regarding this point, it would not be correct to say that in itself a formal or analytic proposition could be proved false or be falsified by new experience or knowledge. What more precisely can occur is that *its domain of application can be restricted or even lost.* For example: since the development of non-Euclidean geometries, the Pythagorean Theorem has lost part of its theoretical domain; it is not the only useful geometry anymore. And since the theory of relativity has shown that physical space is better described as Riemannian, this theorem has lost its monopoly on describing physical space. However, this is not the same as to say that the Pythagorean Theorem has been falsified in a strict sense. This theorem *remains perfectly true* within the theoretical framework of Euclidean geometry, where we can prove it, insofar as we assume the basic rules that constitute this geometry. This remains so, even if Euclidean geometry's domain of application has been theoretically restricted with the rise of non-Euclidean geometries and even if it has lost its full applicability to real physical space after the development of general relativity theory.

The case is different when a law belonging to an empirical science is falsified. In this case, the law *definitely loses its truth together with the theory to which it belongs*, since its truth-value depends solely on its precise empirical application. Newtonian gravitational law, for instance, was falsified by general relativity. It is true that it still has valuable practical applications that do not require the highest level of accuracy. The best one could say in its favor is that it has lost some of its truth, trying to make this idea clear by appealing to multi-valued logic.

13. Conclusion

There is surely much more that can be said about these issues. I believe, however, that the few but central considerations that were offered here were sufficient to convince you that semantic verificationism, far from being a useless hypothesis, comes close to being rehabilitated when investigated with a methodology that does not overlook and therefore does not violate the delicate tissue of our natural language. The fundamental questions of philosophy are as fascinating as difficult because of their subjacent complexity and wideness. Inventing ways to make them easy is to be relieved by illusory answers.

APPENDIX TO CHAPTER V

THE ONLY KEY TO SOLVING THE HUMEAN PROBLEM OF INDUCTION

> It would be impossible to say truly that the universe is a chaos, since if the universe were genuinely chaotic there could not be a language to tell it. A language depends on things and qualities having enough persistence in time to be identified by words and this same persistence is a form of uniformity.
> —*J. Teichman & C. C. Evans*

Here I will first reconstruct in the clearest possible way the essentials of Hume's skeptical argument against the possibility of induction (Hume 1987 Book I, III; 2004 sec. IV, V, VII), viewing it separately from his amalgamated analysis of causality. My aim in doing this is to find a clear argumentative formulation of his argument that allows me to outline what seems to be the only adequate way to react to it in order to re-establish the credibility of inductive reasoning.

1. Formulating a Humean argument

According to Hume, our inductive inferences require support by metaphysical principles of the *uniformity of nature*. Although induction can move not only from the past to the future, but also from the future to the past and from one spatial region to another, for the sake of simplicity I will limit myself here to the first case. A Humean principle of uniformity from the past to the future can be stated as:

PF: The future will resemble the past.

If this principle is true, it ensures the truth of inductive inferences from the past to the future. Consider the following very simple example of an inductive argument justifying the (implicit) introduction of PF as a first premise:

1. The future will resemble the past. (PF)
2. The Sun has always risen in the past.
3. Hence, the Sun will rise tomorrow.

This seems at first glance a natural way to justify the inference according to which if the Sun rose every morning in the past then it will also rise tomorrow, an inference which could be extended as a generalization, 'The Sun will always rise in the future.' We make these inferences because we unconsciously believe that the future will be like the past.

It is at this point that the problem of induction begins to delineate itself. It starts with the observation that the first premise of the argument – a formulation of the principle of the uniformity of nature from the past to the future – is not a truth of reason characterized by the inconsistency of its negation. One could say it is not an analytic thought-content. According to Hume, it is perfectly imaginable that the future could be very different from the past, for instance, that in the future trees could bloom in the depths of winter and snow taste like salt and burn like fire (1748, IV).

We can still try to ground our certainty that the future will resemble the past on the past permanence of uniformities that once belonged to the future, that is, on past futures. This is the inference that at first glance seems to justify PF:

1. Already past futures were always similar to their own pasts.
2. Hence, the future of the present will also resemble its own past.

The problem with this inference is that it is also *inductive*. That is, in order to justify this induction we need to use PF, the principle that the future will resemble the past; but PF itself is the issue. Thus, when we try to justify PF, we need to appeal once more to induction, which will require PF again... Consequently, the above justification is *circular*.

From similar considerations, Hume concluded that induction cannot be rationally justified. The consequences are devastating: there is no rational justification either for expectations created by the laws of empirical science or for our own expectations of everyday life, since both are grounded on induction. We have no reason to believe that the floor will not sink under us when we take our next step.

It is true that we are almost always willing to believe in our inductive inferences. But for Hume, this disposition is only due to our psychological constitution. We are by nature inclined to acquire habits of having inductive expectations. Once we form these expectations, they force us to obey them almost like moths flying towards bright lights. This is an

extremely skeptical conclusion, and it is not without reason that only a few philosophers have accepted Hume's conclusion. Most think that something somewhere must be wrong.

There have been many interesting attempts to solve or dissolve Hume's problem; all of them in some way unsatisfactory.[1] I believe my approach, although only sketched out, has the virtue of being on the right track. I want to first present a general argument and then show how it could influence PF.

2. The basic idea

My basic idea has a mildly Kantian flavor, but without its indigestible synthetic a priori. We can sum it up in the view that *any idea of a world (nature, reality) that we are able to have must be intrinsically open to induction*. I see this as a conceptual truth in the same way as, say, the truth of our view that any imaginary world must in principle be accessible to perceptual experience.

Before explaining it in more detail, I should note that my view is so close to being self-evident that it would be strange if no one had thought of it earlier, as the citation at the start of this appendix proves. More technically, Keith Campbell followed a similar clue in developing a short argument to show the inevitability of applying inductive procedures in any world-circumstances (1974: 80-83). As he noted, in order to experience a world cognitively – as an objectively structured reality – we must continually apply empirical concepts, which, in turn – if we are to postulate, learn from and use them – require a *re-identification* of the *designata* of their applications as identical. However, this is only possible if there is a degree of uniformity in the world that is sufficient to allow these re-identifications. Indeed, if the world were to lose all the regularities implicitly referred to, no concept would be re-applicable and the experience of a world would be impossible.

Coming back to my basic idea, and understanding the concept of world minimally as any set of empirical entities compatible with each other[2], this idea can be unpacked as follows. First, I consider it an indisputable truism

---

[1] For example, Hans Reichenbach (1938), D. C. Williams (1942), P. F Strawson (1952), Max Black (1954), Karl Popper (1959)... Original as they may be, when faced with the real difficulties, all these attempts prove disappointing. (For critical evaluation see W. C. Salmon 1966 and Laurence Bonjour 1998, Ch. 7.)

[2] For the sake of the argument, I am abstracting here the subject of experience... Anyway, this would demand an addition of assumed regularities.

that a world can only be experienced and said to exist if it is at least *conceivable*.[3] However, we cannot conceive of any world without some degree of uniformity or regularity. Now, since we can only experience what we are able to conceive, it follows that we cannot experience any world completely devoid of regularity. This brings us to the point where it seems reasonable to think that the existence of regularity is all that is necessary for at least *some* inductive procedure to be applicable. However, if this is the case, then it is impossible for us to conceive of any world of experience that is not open to induction. Consequently, it must be a conceptual truth that if a world is given to us, then some inductive procedure should be applicable to it.

There is a predictable objection to this idea: why should we assume that we cannot conceive the existence of a *chaotic* world – a world devoid of regularities and therefore closed to induction? In my view, the widespread belief in this possibility has been a deplorable mistake, and I am afraid that David Hume was chiefly responsible for this.[4] His error was to choose causal regularity as the focus of his discussion, strengthening it with carefully selected examples like those of trees blooming in winter and snow burning like fire. This was misleading, and in what follows, I intend to explain why.

Causal regularity is what I would call a form of *diachronic regularity*, that is, one in which a given kind of phenomenon is regularly followed by another kind. We expect the 'becoming' (*werden*) of our world to include regular successions.

However, induction applies not only to diachronic regularities, but also to something that Hume, with his fixation on causality, did not consider, namely, *synchronic regularities*. Synchronic regularities are what we could also call *structures*: states of affairs that endure over time in the constitution of anything we can imagine. The world has not only a 'becoming' (*werden*), but also a 'remaining' (*bleiben*), with its multiple patterns of permanence. And this remaining must also be inductively graspable.

---

[3] After all, conceivability belongs to the grammatical structure of what we understand with the term 'world.' The sentence 'There are worlds that cannot be conceived' is contradictory, for to know the existence of any inconceivable worlds, we must already have conceived them, at least in some vague, abstract sense.

[4] Strangely enough, the idea of a chaotic world to which induction isn't applicable has been uncritically assumed as possible in the literature on the problem, from P. F. Strawson to Wesley C. Salmon. This exposes the weight of tradition as a two-edged sword.

We can make this last view clear by conceiving of a world without any diachronic regularity, also excluding causal regularities. This world would be devoid of change, static, frozen. It still seems that we could properly call it a world, since even a frozen world must have regularities to be conceivable; it must have a structure filled with synchronic regularities. However, insofar as this frozen world is constituted by synchronic regularities, it must be open to induction: we could foresee that its structural regularities would endure for some time – the period of its existence – and this already allows a very strong degree of inductive reasoning!

Considerations like this expose the real weakness in Hume's argument. By concentrating on diachronic patterns and thinking of them as if they were the only regularities that could be inductively treated, it becomes much easier to suppose the possibility of the existence of a world to which induction does not apply or cannot be applicable, a world that nevertheless continues to exist.

To clarify these points, try to imagine a world lacking both synchronic and diachronic regularities. Something close to this can be grasped if we imagine a world made up of irregular, temporary, random repetitions of a single point of light or sound. However, even if the light or sound occurs irregularly, it will have to be repeated at intervals (as long as the world lasts), which demonstrates that it still displays at least the regularity of a randomly intermittent repetition open to recognition. But what if this world didn't have even random repetitions? A momentary flash of light… Then it would not be able to be fixed by experience and consequently to be said to exist. The illusion that it could after all be experienced arises from the fact that we already understand points of light or sounds based on previous experiences.

My conclusion is that a world absolutely deprived of both species of regularity is as such inconceivable, hence inaccessible to experience – a non-world, an anti-world. We cannot conceive of any set of empirical elements without assigning it some kind of static or dynamic structure. But if that's the case, if a world without regularities is unthinkable, whereas the existence of regularities is all we need for some kind of inductive inference to be applicable, then it is impossible that there is for us a world closed to induction. And since the concept of a world is nothing but the concept of a world *for us*, there is no world at all that is closed to induction.

Summarizing the argument: By focusing on causal relationships, Hume invited us to ignore the fact that the world consists of not only diachronic, but also synchronic regularities. If we overlook this point, we are prone to

believe that we could conceive of a world inaccessible to inductive inference. If, by contrast, we take into account both general types of regularity to which induction is applicable, we realize that a world which is entirely unpredictable, chaotic, devoid of any regularity is impossible, because any possible world is conceivable and any conceivable world must contain regularities, which makes it intrinsically open to some form of induction.

One could insist on thinking that at least a world that is chaotic but not entirely chaotic could exist, with a minimum of structure or uniformity, so that it would exist but be insufficient for the application of inductive procedures. However, this is a theoretical impossibility, for *induction has a self-adjusting nature*, that is, its principles are such that they are always conceivably able to be *calibrated* to match any degree of uniformity that is given in its field of application. The requirement of an inductive basis, of repeated and varied inductive attempts, can always be further extended, the greater the improbability of the expected uniformity. Consequently, even a system with a minimum of uniformity requiring a maximum of inductive searching would always end up enabling successful induction.

These general considerations suggest a variety of internal conceptual inferences, such as the following:

> Conceivable cognitive-conceptual experience of a world ↔ applicability of inductive procedures ↔ existence of regularities in the world ↔ existence of a world ↔ conceivable cognitive-conceptual experience of a world…

These phenomena are internally related to each other in order to derive each other at least extensionally, so that their existence already implies these relations. But this means, contrary to what Hume believed, that when properly understood the principles of uniformity should be analytic-conceptual truths, that is, truths of reason applicable in any possible world.

3. Reformulating PF

To show how I would use the just offered proposal to reformulate the principles of uniformity or induction, I will reconsider in some detail PF, the principle that the future will resemble the past. If my suggestion is correct, then it must be possible to turn this principle into an analytic-conceptual truth constituting our only possibilities of conceiving and experiencing the world. – I understand an analytic-conceptual thought-content to be simply one whose truth depends only on the combination of

its semantic constituents; its truth isn't ampliative of our knowledge, in opposition to synthetic propositions, and is such that its denial implies a contradiction or inconsistency (*Cf.* Ch. V, sec. 12).

To show how the aforementioned suggestion could be applied to reformulating the principles of uniformity or induction, it is necessary to reformulate PF. If my general thesis is correct, then it must be possible to turn this principle into an analytic-conceptual truth constituting a way of conceiving and experiencing the world. Here is a first attempt to reformulate PF in a clearly analytic form:

PF*: The future must have *some* resemblance to its past.

Unlike PF, PF* can easily be accepted as expressing an analytic-conceptual truth, for PF* can be clearly seen as satisfying the above characterization of analyticity. Certainly, it belongs to the concept of the future that it is the future of its own past. It cannot be the future of another past belonging to some alien world. If a future had nothing to do with its past, we could not even recognize it as being the future of its own past, because it could be the future of anything, what seems incoherent. In still clearer words: the future of our actual world *W*, as F*W*, can only be the future of the past of *W*, that is, P*W*. It cannot be the future of infinitely many possible worlds, *W*1, *W*2, *W*3... that have as their past respectively P*W*1, P*W*2, P*W*3... Thus, it is necessary that there must be something that identifies F*W* as being the future of P*W*, and this something can only be some degree of resemblance in the transition.

Against this proposal, we can try to illustrate by means of examples the possibility of complete changes in the world, only to see that we will always be unsuccessful. Suppose we try to imagine a future totally different from its past, a 'complete transformation of the world' as described in the *Book of Revelations*. It is hard to imagine changes more dramatic than those described by St. John, since he intends to describe the end of the world as we know it. Here is the biblical passage describing the locusts sent by the fifth angel:

> In appearance the locusts were like horses equipped for battle. And on their heads were what looked like golden crowns; their faces were like human faces and their hair like women's hair; they had teeth like lions' teeth and they wore breastplates like iron; the sound of their wings was like the noise of horses and chariots rushing to battle; they had tails

> like scorpions with stings in them, and in their stings lay their power to plague mankind for five months.[5]

At first glance, these changes are formidable. Nonetheless, there is nothing in this report that puts PF* at risk. In fact, closer reflection on the example demonstrates that even PF isn't seriously challenged. Although these biblical locusts are indeed very strange creatures, they are described as combinations of things already familiar to us. These things are horses, women, hairs, men, heads, teeth, scorpions' tails with stings, human faces, etc. Both internally and externally, they include a vast quantity of synchronic regularities, of permanent structural associations, together with familiar diachronic associations, like the causal relationship between the noise produced and the movement of wings or the sting of the scorpion and the effects of its poison on humans…

In fact, were it not for these uniformities, the apocalypse as described by St. John would not be conceivable, understandable and able to be the subject of any linguistic description. The future, at least in proportion to its greater proximity to the present, must maintain sufficient similarity to its past to allow an application of inductive procedures to recognize the continuity of the same world we know today.

Now one could object that maybe it is possible that at some time in a remote future we could find a dissimilarity so great between the future and our past that it invalidates any of our reasonably applicable inductive procedures – a remote future that would be radically different from its past. Indeed, it seems conceivable that a continuous sequence of small changes could in the course of a very long period of time lead to something, if not completely different, at least extremely different. Nevertheless, this would not discredit PF*, because its formulation is too weak, requiring only that *some* similarity must remain. However, it also seems that this weakness of PF*, even if not robbing it of its analytic-conceptual character, exposes PF* to charges of disproportionate poverty as a way to assure the reliability of our inductive projections.

However, precisely this weakness of PF* suggests a way to improve it. It leads us to see that the closer we get to the point of junction between the future and the past, the greater must be the similarity between future and past, both becoming identical at their limit, which is the present. We can approximate this issue by remembering the Aristotelian analysis of *change* as always assuming the permanence of something that remains identical in a continuous way, without gains or losses (Aristotle 1984, vol 1: *Physics*,

[5] *Revelation of St. John* 9, 7.

200b, 33-35); in other words, the intuitive idea is that every change must occur on some basis of permanence.

This leads us to create another variant of PF, namely, the principle according to which in a process of change the amount of permanence must be inversely proportional to the period of time in which the change occurs. In other words: if there is a sequence of changes that are parts of a more comprehensive change, the changes that belong to a shorter sequence typically presuppose a greater number of permanent structural (and sequential) associations than the sequence constitutive of the more comprehensive change.

This principle can be illustrated with numerous examples. Consider a simple one: the changes resulting from heating a piece of wax. The change from the solid state to the liquid state presupposes the permanence of the same wax-like material. However, the next change, from liquid wax to carbon ash, presupposes only the permanence of carbon atoms. If the heat then becomes much more intense, carbon will lose its atomic structure, giving place to a super-heated plasma of subatomic particles. We have here a sequence of four time periods: regarding the shortest period of time from t1 to t2, we assume that we will be left with (i) the same wax, made up of (ii) the carbon molecules and atoms, which in turn are composed of (iii) their same subatomic constituents. In the longer period of time from t1 to t3 we assume the identity of only (ii) and (iii): carbon atoms and subatomic particles. And in the still longer period of time from t1 to t4 the only things that remain the same are (iii): subatomic constituents.

Note that this model is not restricted to changes in the physical material world! As Leibniz saw: *Natura non facit saltus*. The same examples repeat in every domain that one can imagine, chemical, biological, psychological, social, economic, historical… with the same patterns: the closer the future is to its junction with its past, the more structural identities must be in some way assumed. For example: the process of industrialization. The Industrial Revolution was a period of social and economic changes from an agrarian society to an industrialized society, which suffered an upheaval in the mid-19th century. As a whole, after its second period it included the refinement of the steam engine, invention of the internal combustion engine, harnessing of electricity, construction of infrastructure such as railways… and, socially, the more complete exodus of families from rural areas to large cities where factories were constructed… However, when we choose to consider a short period in this process, for instance, at the end of the 18th century, the only outstanding changes were probably the invention of a simple piston engine and a minor exodus from

the countryside, most characteristics of society otherwise remaining essentially the same.[6]

We conclude that it is intrinsic to the very structure of the world of experience – and of *possible* experience – that changes taking place in a shorter period of time tend to presuppose more permanence than the most comprehensive long-term changes within whose course they occur. Consequently, the future closer to its present should as a rule inevitably be more similar to its past in more aspects than more distant future will be (as already noted, the far distant future may be almost unrecognizably different from the present). At the point of junction between future and past (the present), no difference will be available.

Regarding induction, this principle assures that inductive predictions will become more likely the closer the future is to the present. On this basis, we can improve the principle PF* as:

> PF**: As a rule, the closer the future is to the junction point with its own past, the more it will tend to resemble its past, the two being indistinguishable at the point of junction (the present).

For a correct understanding of PF**, we must add two specifying sub-conditions:

(i) that this principle should be applied to a future that is sufficiently close to its past and not to an indefinitely distant future.
(ii) to safeguard the possibility of anomalous but conceivable cases in which we find shorter sequential periods where states of affairs of a more distant future are closer to the present than those of the near future.

Although I admit that PF** deserves more detailed and precise consideration, it seems to me intuitively obvious that so understood this principle already meets a reasonable standard of analyticity.

Moreover, it is the truth of PF** which explains why it is natural for us to think that the more distant the future, the less probable our inductive forecasts will be. This is the very familiar case of weather forecasts: they are presently reliable for two or three days, less so for a week or more... It also explains why our inductive generalizations about the future cannot be

[6] One could still object with cases like that of someone who suddenly awakes from a dream… But one forgets the remaining fact that it is the very same person who was dreaming that is now awakened.

applied to a very distant future. For instance, through induction we can infer that the Sun will 'always' rise, but *always* must be placed in quotation marks. On the basis of induction, it makes sense to affirm that the sun will rise tomorrow morning or even a thousand years from now. But it defies common sense (and is for cosmological reasons false) to use the same inductive basis to claim that the Sun will still rise every morning in seventeen billion years.

How PF** applies is circumstantially determined. If the future is sufficiently close to its junction with the past, then the future will be unavoidably similar to its past. The problem, of course, is that we need to establish criteria for judging how close in time the future must be to its past so that PF will still apply. We can speculate as to whether the answer does not depend on the background represented by the domain of regularities in which we are considering the change – a domain of regularities to which a whole system of sufficiently well-entrenched beliefs applies.

For example: the inductive conclusion that the Sun will rise tomorrow belongs to a domain of regularities that may someday undergo changes predicted by contemporary cosmology. This may include a very distant future in which dramatic changes, such as the death of the Sun, are also predictable based on the astronomically observed fates of similar stars in our universe.

Of course, it is always possible that the Sun will not rise tomorrow! However, this is only conceivable at the price of an immense loss of other well-entrenched beliefs about astronomical regularities and, subsequently, the loss of the current intelligibility of a considerable portion of the physical world around us. Still, what makes us consider as highly likely the future occurrence of regularities such as that the Sun will rise tomorrow?

The ultimate answer seems to be based on the inevitable assumption that our world will continue to exist as a system of regularities, at least in the form prescribed by PF**. However, this assumption seems to be a blind *gamble*! After all, there is nothing preventing our whole world from suddenly disappearing. However, the impression of a paradox evaporates as soon as we consider that this hypothesis is completely unverifiable. If our whole world suddenly disappear and there is no other, how can we know this after we have also disappeared with the world? Now, if the hypothesis is unverifiable, it must be senseless.[7] In contrast, the hypothesis

[7] In its lack of sense, the question remembers the anthropic principle. The question, 'Why is possible that we are able to think the world?' loses its sense as soon as we

that our world will continue to exist can be verified in the future, hence it is meaningful. Because of this asymmetry, we are free to accept that since we cannot really think that there will be no future at all, the regularities of our world will need to take the form prescribed by PF**, that is, we are inevitably led to admit that certain domains of cohesive regularities will have some permanence.

The above outlined argument concerns just a single form of induction: from the past to the future. Nevertheless, the attempt to better specify it and to generalize about further developments would be worthwhile, since it suggests a path free of insurmountable hindrances. This may be of some interest regarding a problem that from any other angle seems to remain disorienting and intangible.

---

consider that under infinitely many possible worlds this is one under the few able to produce conscious beings able to pose this pseudo-question.

CHAPTER VI

SKETCH OF A UNIFIED THEORY OF TRUTH

Das wahre Bild des Fehlers ist das indirekte Bild der Wahrheit; das wahre Bild der Wahrheit ist der einzig wahre.
[The true picture of the error is the indirect picture of truth; the true picture of truth is the only true one.]
—Novalis

He who thinks the separated to be separated and the combined to be combined has the truth, while he whose thought is in a state contrary to that of the objects is in error.
—Aristotle

We have drawn some conclusions from the previous chapters: the cognitive meaning of an assertoric sentence is a semantic-cognitive rule, namely, its verifiability rule, which is the same as an e-thought-rule, a spatio-temporally extensible proposition in the explained sense – the primary truth-bearer. The verifier of a proposition is the fact it represents, a complex entity constituted of tropical arrangements. Moreover, consistent with our idea that the effective applicability of a (possible) conceptual rule in its domain is the same as the *existence* of a trope or a cluster of tropes able to satisfy it, we can expect that by symmetry the effective applicability of a verifiability e-thought rule in its proper context should be the same as the *existence of the fact* that satisfies it. Finally, since the property of a verifiability e-thought rule of being effectively applicable was devised as the reason we call it *true*, it seems that the existence of the fact it refers to should be the same as its truth. This is a strange conclusion.

Moreover, this conclusion seems at odds with another view, namely, the correspondence or adequation theory of truth, according to which the truth of an e-thought-rule (of a proposition) is its correspondence with a fact and not the existence of the fact referred to by it. This is somewhat disturbing, for as already noted we have the best methodological reasons for defending truth as correspondence. This theory expresses a modest (even lexicalized) commonsensical view with a long tradition.

Historically, it has been the standard truth-theory from Plato to the nineteenth century, and even nowadays most theorists are inclined to accept it. Notwithstanding, existence and truth, as, respectively, the effective applicability of a verifiability rule and the correspondence with a fact, do not seem to have much in common.

Nonetheless, I believe to have found a way to overcome the difficulty. The solution consists in remembering that, as dictionaries show, the word 'truth' has two very distinct main bearers in natural language (*Cf.* Ch. IV, sec. 31). Indeed, among a variety of irrelevant senses, dictionaries almost always distinguish clearly between two common attributions of truth:

(a) *thought-truth*, which is the 'truth of a thought in conformity with things being as we believe they are,' (One could say, the *property of an e-thought-rule* of being satisfied *by a corresponding fact*), and
(b) *fact-truth*, truth as the '*actual*, *real or existing thing or fact*.'

Even if thought-truth is primary and fact-truth derivative, my suggestion was that fact-truth is more properly identified with *existence* – the existence or reality of a fact, which is the same as the *dispositional* higher-order property of a fact of having its own verifiability rule effectively applicable to itself (both things, the fact and its higher-order property being simultaneously given). On the other hand, truth attribution in the archetypical sense of thought-truth continues to be reserved to the metaproperty of the actual e-thought/verifiability rule of being effectively applicable to a fact. In this case, we see the effective applicability as the correspondence or adequacy with a fact, which as a property of the verifiability rule must also be a higher-order property-trope regarding the fact to which the rule is applicable.

It is important to see that although thought-truth and fact-truth might at first glance seem to be only two different ways to consider exactly the same thing, there is a fundamental difference between them. Thought-truth is the truth of an e-thought-content-rule that is considered effectively applicable to its corresponding fact. This attribution of truth to an e-thought requires the verifiability rule constitutive of the e-thought to be effectively applicable to its fact, which implies the *real existences* of (i) the fact, (ii) the e-thought as a verifiability rule, and (iii) at least one cognitive being who has reasons to be aware of its applicability, while in many cases the reason is simply its application by him. However, fact-truth – the reality or existence of the fact – demands much less. It demands only (i): the existence of the fact, understood as the dispositional trait of being able to have a *possible* verifiability e-thought rule effectively

applicable to it. It does not demand either the actual existence of the verifiability rule or the existence of an epistemic subject able to apply this rule to it! In a world without cognitive beings, these rules and their applicabilities would be purely *dispositional* tropical properties in the sense that if there were cognitive beings able to know the facts, they could construct these e-thought-rules and effectively apply them. As we have already realized (Ch. IV, sec. 35), a world without cognitive beings would have fact-truths but no thought-truths.

Ernst Tugendhat was right in holding that correspondence and verifiability cannot be separated (1983: 235-6) and we can now see why. It must be so because in considering the verifiability/non-verifiability of an e-thought-rule we need to find a corresponding match/mismatch between the dependent criterial configurations demanded by the verifiability e-thought rule and the corresponding contingent arrangement of tropes called the real fact that satisfies or does not satisfy this demand by either having or not having the independent criterial configurations. And this match, even if first concerning sub-facts, must at least indirectly concern the grounding fact, since the former are only aspects or facets of the latter. (See Ch. IV, sec. 25-27).

Based on what we have learned thus far, the purpose of this last chapter is to outline a correspondence analysis of truth in sufficient detail to make it more complete and plausible than what we have seen in philosophy until now – an analysis with the potential not only to better clarify the distinctions we have made, but also an attempt to take some account of the problem in its real complexity.

1. Deceptive simplicity of correspondence

I begin by addressing the shallowest objection against the correspondence theory of truth. It is the claim that the theory is nothing but a trivial, empty truism. According to this widespread objection, to say that truth is agreement with facts is a too obvious platitude to deserve philosophical attention (Blackburn 1984, Ch. 7.1; Davidson 1969).

The illusion that feeds this objection emerges from the fact that all too often in philosophy careful scrutiny has shown that what initially seems to be a plain, uncomplicated meaning conceals unexpected complexities. One impressive example of this was the causal theory of action. Who could at first glance foresee that analysis would show that such an apparently simple thing as the concept of human action could involve a variety of sometimes very complex processes, like the formation of reasons (made from desires and beliefs) producing previous intentions that at the right

time produce the intention-in-action (the *trying*) directly causing the right bodily movements, which should produce as final outcome the intended effects? In what follows, I hope to convince you that the correspondence theory of truth is no exception to this rule. The supposed simplicity of the correspondence relation is only apparent, revealing our lack of awareness of what we really do when making truth-claims.

Methodologically, my strategy consists in reconsidering the best insights that we have inherited on the correspondence theory and in asking how far they can be developed and plausibly combined in order to lead us to a full-blooded philosophical analysis of the correspondence relation. As you will see, this endeavor ultimately requires a pragmatic investigation of the dynamic constitution of correspondence, which in the end exposes its intrinsic relationship with verifiability, coherence, criteria of truth and even its dependence on an adequate answer to the problem of perception.

Consequently, in order to bring clarity to our views, what we need is to delve more deeply into the waters of the above suggested approach to the correspondence theory of truth.

2. Analysis of correspondence (1): structural isomorphism

Suppose that truth in a privileged sense is indeed correspondence (adequation, agreement, match…) between a verifiability e-thought-content rule and the fact it represents. In this case, we must first specify each term of this definition. We have already clarified the concept of thought as an e-thought – an extensible thought-content properly built upon psychological p-thought-rules, as the archetypical truth-bearer in our discussion of Frege's semantics (Ch. IV, sec. 34). We did this along with a detailed defense of the idea that an elementary real fact is a cognitively independent arrangement of elements, which are tropical properties and clusters of compresent tropes corresponding to a proper singular statement. And as we also saw, 'fact' is an umbrella-term that includes actual static facts (situations, states of affairs…) and dynamic facts (events, processes…), serving in this way as universal truth-makers – the most proper verifiers of statements (Ch. IV, sec. 23). What is now in need of analysis is the concept of correspondence in its relevant sense.

The early Wittgenstein, as is well known, insightfully defended a correspondence or adequation theory of truth in the form of a pictorial theory of representation in his *Tractatus Logico-Philosophicus* (1984g, sec. 2.21). I prefer to ignore the implausible atomistic metaphysics of this work, later rejected by him, though not its deeper insights (*Cf.* Stenius:

1981[1]); and one profound insight of the *Tractatus* is that a fundamental condition of representation is a pictorial relationship between the logically analyzed sentence (*Satz*), expressing what he calls a *thought* (*Gedanke*),[2] and the *possible fact*, called a state of affairs (*Sachverhalt*), which can be actualized as the *real fact* (called by him a *Tatsache*[3]), the fundamental verifier able to make the thought true. The idea was resourcefully explored by E. G. Stenius in his important monograph on the *Tractatus* (1960) and several later articles (particularly that of 1981) by applying to it the mathematical concept of *structural isomorphism*.

Applied to the correspondence view of truth, a true thought-content must have at least structural isomorphism with a possible (conceivable) or actual (real) fact. As I understand it, the structural isomorphism is constituted by three conditions, which are at least partially explanatory of the idea of correspondence:

(i) A *bi-univocal relation:* each semantic component of a verifiability e-thought-content rule (or of the sentence adequately expressing it) and each corresponding element constituting the *possible* or *actual fact* (understood as something epistemically objective[4]) must have a biunivocal relation.
(ii) A *concatenation*: the component rules of a verifiability e-thought-content rule (or of an analyzed sentence) must be combined in the *same manner* as are the elements composing the possible or actual fact.
(iii) A *correlation:* a verifiability e-thought-content rule as a whole must be biunivocally related to the possible or actual (real) fact,

---

[1] E. G. Stenius suggested that although rejecting the logical atomism of the *Tractatus*, Wittgenstein withholds the pictorial view in his later book, *Philosophical Investigations*, what is shown in his suggestion that the photo of a boxer in a particular stance is a *sentence-radical* (*Satzradical*) susceptible to different interpretations (like the varied possible illocutionary forces of an utterance) (*Cf.* 1984c, sec. 22, note). The sentence-radical must possibly express an e-thought representing a real person in a pictorial way.

[2] In a letter to Russell from 1919 Wittgenstein explained that for him a thought consisted of *psychical* elements (Wittgenstein 1974). This liberates him from Frege's Platonist commitment.

[3] I follow here Stenius' interpretation of the *Sachverhalt/Tatsache* distinction (1964: 31).

[4] I will use the phrase 'actual fact' in the sense of an epistemically objective real fact, understanding the word 'objective' in John Searle's sense of *epistemically objective*, which includes not only external but also internal (psychological) facts, insofar as they are intersubjectively sharable. (*Cf.* chapter IV, sec. 34)

making it its *correlate*. (This condition presupposes the satisfaction of (i) and (ii), though it goes beyond them.)

Now, a necessary condition for the truth of an e-thought-content-rule is that it must be structurally isomorphic with an *actual* fact in the world. And a necessary condition for its falsity is that although it is structurally isomorphic with a *possible* fact – that is, a conceivable or imaginable fact – it is not structurally isomorphic with any expected contrastively real fact. Structural isomorphism is a necessary condition, because a supposed proposition, an e-thought-content-rule, must be at least possibly classifiable as true or false in order to have any cognitive function and deserve its name.

Notice that we do not need to believe that possible facts exist in some Platonic realm in order to accept the requirement of conceivability. To conceive or imagine a possible fact is simply a *psychological phenomenon,* and it is clear that we don't need to conceive or imagine it in all the details we would be forced to consider if it were a real-actual fact. In other words, a true verifiability e-thought rule must be isomorphic with a fact in the world, while a false e-thought-rule, though not isomorphic with a fact in the world, must at least be isomorphic with a possible (conceivable, imaginable) fact, because it is by means of this 'projection' that we know that in principle it could be correlated with an actual fact in the world.

The natural way to apply this view to real statements is to begin with singular predicative or relational statements in their actual linguistic practices, taking the logically analyzed sense-components of their sentences as the elements that must be biunivocally related to the elements of the possible or actual facts. Thus, we begin with e-thought-contents expressed by singular statements of the form *Fa* (ex.: 'John is easygoing') or *aRb* (ex.: 'John is the father of Mary') or *Rabc* (ex.: 'John gives Mary a flower') or *Rabcd* (ex.: 'John gives Mary a flower to please Jane')… In order to be *true*, these statements must at least satisfy the same already explained conditions of structural isomorphism:

(i) Each component sense or semantic-cognitive rule expressed by each nominative and predicative expression must correspond *biunivocally* to the respective elements constitutive of the respective fact in the world. This fact is an arrangement made up of simple or complex property-tropes (like being easygoing, being the father of, giving something to someone, giving something to someone to please someone else) and tropical objects, made up at least of tropes like those of form, solidity, mass… and possibly

mental states like feelings… all of them displaying compresence (John, Mary, Jane, the flower…) (See Ch. IV, sec. 5)

(ii) The *concatenation*, i.e., what we called a *manner of connection* between the component rules of the e-thought-rule and the fact, must be preserved. Because of the manner of connection, a sentence with the form *Fa* cannot be replaced by *aF* (e.g., 'John is easygoing' cannot be replaced by 'Easygoing is John'), and a sentence with the form *aRb* cannot represent the fact represented by *bRa* (e.g., 'John is the father of Mary' cannot be replaced by 'Mary is the father of John'). Regarding these forms and orders of connection, they emphasize that properties are relatively dependent on objects in the context of the facts they belong to (being easygoing depends on John's existence, being a father depends on the existence of John and Mary…); the concatenation can be already read in the components. [5] (See Ch. IV, sec. 7-9)

(iii) the *whole* thought-content must be biunivocally related with its possible or actual corresponding fact.

This view should apply even to complex and vague expressions. Take, for instance, statements like 'Céline had a strange personality' and 'The Irish potato famine was caused by late blight.' Insofar as these expressed thought-content-rules can be objectively-interpersonally verified, they are acceptable. Although it is surely not so easy to explain Céline's strange personality or how the late blight caused the Irish potato famine, these concepts remain open to investigation and reducible to complex associations of tropes.

3. Analysis of correspondence (2): categorial match

According to Stenius, sharing the same ordered logical structure isn't enough. He was aware of this difficulty when he suggested that there must be what we could classify as a condition (iv), demanding some kind of *categorial match* between each biunivocally related pair of elements. Using his own words, we could say that the components of the e-thought-rule must be *indices* of the elements of the fact they biunivocally represent.

Since words like 'categorial match' and 'indices' are not very informative, one could search for something less metaphorical. As I have

[5] The analysis could certainly not go further, requiring that there must be some *R1* relating *F* with *a* in *Fa*, etc. (*Cf.* Appendix to Chapter III, sec. 1)

already noted (Ch. IV, sec. 3), Kant wrote about *schemata*. For him, a concept is a rule able to produce *figure-types* or *patterns* (*Gestalten*) that we can *correlate* with the objectively given in order to recognize it. As he wrote:

> The concept of dog means a rule according to which my imagination in general delineates the figure [pattern] of a four footed animal, without being limited to any particular figure offered by experience or by any possible image that I can represent *in concreto*. (1988, A 141)

Although Kant's full exposition of this topic is frustratingly obscure, it seems clear that it anticipates what we have previously learned in our readings of Wittgenstein, Frege and later philosophers (particularly Michael Dummett and Ernst Tugendhat), suggesting that we look for an answer in terms of the specifying power of semantic-cognitive rules. Restricting ourselves to the simplest case of the singular predicative statement, what we have is the following. First, consider the conceptual senses expressed by singular and general terms, namely, the identifying and ascription rules along with their joint formation of a verifiability rule. Each of these semantic-cognitive rules is able to establish an undetermined variety of dependent criterial configurations, whose satisfaction is nothing but their matching with independent or external criterial tropes or configurations of tropes (properties), clusters of selected compresent tropes (objects), arrangements of such configurations of tropes and such clusters (facts). Once all these dependent criterial configurations are seen as satisfied by suitable tropical arrangements or actual facts in the appropriate context, the verifiability rule is considered effectively applicable. Since this rule is nothing but the e-thought, once definitely applicable this verifiability e-thought-content rule will be called true and said to represent a fact. This shows that Stenius' indices, Kant's schematized patterns, and our Wittgensteinian criteria or criterial configurations are only increasingly detailed attempts to do the same thing, namely, to isolate, to distinguish in their uniqueness the isomorphic elements constitutive of the represented facts, in order to justify the applicability of their verifiability rules.

Since these semantic-cognitive rules are also senses in a Fregean conception of 'modes of presentation,' what we first need to add to our understanding of correspondence as structural isomorphism are the *individualizing senses* of the component expressions, that is, the semantic-cognitive criterial rules constitutive of the verifiability e-thought rule. As explained in Chapter IV, we typically identify the grounding fact corresponding to the basal e-thought by means of some variable criterial

aspect: a sub-fact. In order to achieve this, what we usually do is the following. By means of the partial structural isomorphism between the criteria demanded by the derived verifiability e-thought rules and sub-facts as independent criteria, we usually infer the isomorphism between the cognitive rules constitutive of the basal e-thought (e.g., an identification rule and an ascription rule building the verifiability rule) and the grounding fact, if the e-thought-rule is true, or merely a conceivable grounding fact, if the e-thought-rule is false.

Furthermore, we must remember that we can make any of these rules explicit by means of *definitions able to bring their criteria of application to the surface*, as I have initially shown using the concept of chair as an example (Ch. II, sec. 7). In the aforementioned examples, we can do something similar. Concerning names, in examples like (1) 'The book is on the table,' (2) 'Kitty is in the kitchen,' and (3) 'John is father of Mary,' this would be done by means of the (semantic-cognitive) criterial definitions given by the identification rules of the nominal terms 'the book,' 'the table,' 'Kitty,' 'the kitchen,' 'John,' 'Mary' (See Appendix Ch. I). Concerning predicative expressions in examples (1), (2) and (3) this would be done by means of definitions of the relational predicative expressions '…is on…', '…is in the…', '…is the father of…' Such definitions will also show how the elements can or cannot be adequately concatenated one with another (the table cannot be on the book, the kitchen cannot be in Kitty, Mary cannot be the father of John) and, mainly, how they can be applied to the tropical elements constitutive of the corresponding grounding fact. The condition (iv) of categorial match is also the condition that *the e-thought and its elements can be made explicit as semantic-criterial rules able to match their proper references*, distinguishing each statement by its proper semantic content.

These explanations entitle us to suggest that when two e-thoughts *p* and *q* display structural isomorphism and the (semantic-cognitive) criterial rules that form the required elements of *p* are the same as the (semantic-cognitive) criterial rules that form the required elements of *q*, then both e-thoughts are at least qualitatively the same. That is, since the cognitive verifiability rules are the same, *p* and *q* express what we may call the same senses, the same verifiability e-thought rules or, as we can also say, the same *contents*. We will return to this point later, when we arrive at the pragmatics of the correspondence relation.

4. Analysis of correspondence (3): intentionality and causality

There are two additional elements that we need to consider in order to complete our analysis of correspondence: (v) *intentionality* and (vi) *causality*. In judging something to be true, we must be aware that we are applying a verifiability rule to a fact, we need to have *a referential directionality* that leads us from semantic-cognitive rules to the tropical criteria that should satisfy them, from an e-thought to the real-actual fact it aims to represent in judgment.[6] One could say that intentionality gives to the correspondence a 'mind-to-world direction of fit,' defining the direction of fit of a mental state as what John Searle would call its 'responsibility' to fit an independently existing reality, that is, as its 'mind-to-world responsibility of fit' (2004: 167-9). Intentionality isn't a condition on the same level as the others. It is added to the above explained conditions of correspondence as something belonging to the broader structure of our consciousness, since it requires some kind of conscious attention on the speaker's side. The upshot is that correspondence restricted to isomorphism is symmetrical, while correspondence cum intentionality is asymmetrical. Since correspondence (as agreement, accordance, congruence, conformity, matching) is symmetrical (if A corresponds to B, then B must correspond to A) – correspondence can be the best word for distinguishing isomorphism, but not as well isomorphism regarding categorical match and still less isomorphism regarding intentionality. Less neutral words would be 'picturing,' 'adjustment' and 'adequation,' since these relations are less forcefully symmetrical (if A pictures B, B does not literally picture A, if A is adequate to B, B isn't necessarily adequate to A…). Because of this I will give preference to the word 'adequation,' meaning by it correspondence cum intentionality.

Finally, from the opposite direction, what we may find in the case of true e-thoughts will be (vi) *a suitable causal relation* by means of which an actual fact may make us recognize the truth of its e-thought. However, it seems that in the real world the causal network is so extremely complex

[6] We can intentionally produce factual contents that we call true by acting in the world in order to change it in accordance with our views. For this reason, constructivist philosophers followed Giambattista Vico in the attempt to reverse the direction of fit of the correspondence: we are the ultimate truthmakers. However, this is not correct. Even in this case the truthmaker of the proposition, as the product of human action, is the final fact in the world and not the idea that has produced human action. That is: we can make the fact that makes the truth, but not the fact as truth.

that we are led to see that the causal relation we are considering in no way needs to be a direct one. In fact, it can be very and even extremely indirect, easily misleading us to the belief that it does not exist.

Causality has a 'world-to-mind direction of fit' or a 'world-to-mind responsibility of fit' in the sense that it is what causes thought-content to match reality. We can speak here of the effective applicability of its verifiability procedure and, in the case of some unintentional perceptual indexical e-thought-rule, even of the causal transient construction of such a rule in a given context – a rule that once constructed is intentionally applied.

5. Exemplifying adequation

For reasons of clarity, I will consider a final example of a composite thought-content adapted from Stenius, which allows us to summarize what we have learned about adequation. If someone says: 'John (j) is the father (F) of Peter (p) and of Mary (m), who is a violinist (V)', the logical structure of the e-thought-content expressed by the statement is:

1. jFp & jFm & Vm.

Assuming that we know the identification rules for John, Peter, and Mary, and with the ascription rules of the predicates '…is the father of…' and '…is a violinist,' along with the semantic rule of application of the logical operator '&' (which can be provided by a truth-table), we know that this statement might be true. In other words, we know that we can combine these semantic-cognitive rules, applying them imaginatively in order to conceive a possible state of affairs corresponding to the e-thought-content, giving to the statement at least a meaning. If the statement is false, the correspondence stops here, as a mere adequation with a possible but non-actual fact. Now, suppose that statement (1) is true. In this case, we have the five conditions of correspondence satisfied by a complex real fact. The satisfaction of these five conditions can be presented at least as follows:

(i) The bi-univocal relation between each of the non-logical and logical components rules of the composed verifying e-thought-content rule expressed by (1) and each corresponding factual tropical element.
(ii) The same concatenation between the semantic-cognitive rules constituting the verifiability rule of each singular e-thought-content, including the relations of conjunction among them and

among the biunivocally related elements of the three represented elementary facts.

(iii) The bi-univocal relation between each singular verifiability e-thought-content rule and its represented elementary fact (the same regarding the composed e-thought-rule and composed fact).

(iv) Concerning each verifiability e-thought-content rule, the matching (or satisfaction) of the dependent criteria formed by each component semantic-cognitive rule (identification plus ascription rule) regarding its proper objective correlate – a tropical criterial correlate – together with the two rules of conjunction, assuring us the proper individuation of the rules working as meaning unities.

(v) The intentionality (directionality) we link to the rules, leading us to distinguish what is representing – a composite e-thought-content rule – from what is being represented – the actual corresponding composite fact – building a mind-to-world direction of fit.[7]

(vi) The assumption that 'jFp & jFm & Wm' is true because we have reasons to believe that it is, even if in some very indirect way, causally determined by the facts, the causation having a world-to-mind direction of fit.

As we know, for a disjunctive statement with the form 'jFp ∨ jFm ∨ Vm' to be true, one of the disjuncts, at least, must represent not only a possible fact, but also the *actual* fact, by having criterial configurations of its verifiability rule matched by the external tropical structure of its represented fact. Finally, all false disjuncts must at least correspond to possible (conceivable or imaginable) facts, if we want the statement as a whole to remain cognitively meaningful.

6. Compatibility between verification and correspondence

Against the correspondence view of truth, there is also the objection that it is incompatible with verificationism. The objection can be as follows: a statement can be verified *in many different ways*, insofar as its verifiability rule may be satisfied by an indeterminate range of diversified sub-facts, which are tropical arrangements acting as verifiers. By contrast, correspondence should be a one-to-one relation: the fact corresponding to a true proposition should be univocally related to the proposition stated by

[7] In my view, the addition of intentionality, the sense determines (*bestimmt*) the reference in Frege's way of speaking, or turns into a meaning-fulfilling intention (*Bedeutungserfüllende Intention*) in Husserl's way of speaking.

the assertoric sentence. Consequently, it does not seem possible that what verifies the stated proposition is a corresponding fact, as claimed by traditional correspondence theory. (e.g., Hallett 1988: 29)

As you have probably already noted, the above argument is deeply misleading and it only reaches its conclusion by searching for correspondence in the wrong place. Usually, correspondence requires more than a match between an e-thought-rule and a fact in the world. The verifiability e-thought rule might accept multiple correlative independent criterial tropical arrangements for its application, as it was already clear as we distinguished correspondence with sub-facts from correspondence with the grounding fact (Ch. IV, sec. 26) and later, as we considered Wittgenstein's examples (See Ch. V, sec. 3, his example of the Cambridge boat race).

Correspondence is here, we can say, often a relation between an immediately-derived verifiability e-thought rule and its sub-fact belonging to a more encompassing relation between a mediated-basal verifiability e-thought rule and its grounding fact. And what interests us most here is the last case: the correspondence between the basal verifiability e-thought rule as a typically ramified verifiability rule and the grounding fact with its many different sub-factual manifestations. Concerning this, the central point is that our resulting awareness of the grounding-fact can be *inferred* from the satisfaction of this or that isomorphic sub-fact, viewed as a partial independent external criterial tropical arrangement, often the only one immediately experienced. For instance: I say that I see the grounding fact that there is a ship in the bay, even though I can see this grounding fact only from one side and at a certain distance, that is, by visualizing the specific tropical arrangement constitutive of a sub-fact I already have enough criteria for the inference of the grounding fact. Hence, the comprehension of a grounding fact is typically indirect and inferential. Summarizing, correspondence often occur on two different levels:

1) *Immediate-derived level*: as the matching between the dependent criterial configuration generated by some derived verifiability e-thought rule and the independent criterial configuration formed by some tropical arrangement constitutive of the appropriated sub-fact (e.g., 'I see the side of a ship').
2) *Mediated-basal level*: as the match between a basal verifiability e-thought rule with its many ramifications (i.e., encompassing a great variety of probably true albeit non-verified e-thoughts) and the grounding fact to which the many tropical arrangements constitutive of its derived sub-facts very probably belong. Normally

> it is the satisfaction of a suitable criterial configuration by means of a sub-fact that enables us to indirectly infer[8] the correspondence between the basal e-thought-rule and the grounding fact (e.g., 'I see a ship because I see the side of the ship').

These two levels of correspondence work in the same way, requiring what can be called a two level structural isomorphism, first between the derived though and the sub-fact, second between the basal thought and the grounding fact.

In more detail, the intrinsic relation between verification and correspondence can be explained as follows. First, concerning the immediate-derived level of correspondence, we have a derived e-thought constituted by its own verifiability rule that can be expressed by a sentence of the form *Fa*, for instance,

(a) A ship-bow is rising,

when a ship is approaching the pier. This rule is satisfied by the sub-fact that a ship-bow is rising, which requires the satisfaction of the criteria for an identification rule for the singular term 'a ship-bow there,' added by the satisfaction of the criteria for application of the ascription rule of the predicate '…is rising.' Clearly, the relation between this derived e-thought-rule and the criterial configurations requires structural isomorphism with the independent external criterial tropical arrangements. But in this case and in many others, the immediate-derived level of correspondence can ground the inference of a mediated-basal level of correspondence. In this case, by means of the experience of sub-facts as independent criteria, we may indirectly infer that overall a main verifiability e-thought rule with its wider divisions corresponds to a whole grounding fact. Consider for instance the e-thought-rule expressed by the following statement of the form *Fa*: 'That ship is approaching.' We may conclude the truth of this statement simply by means of the already verified criterial dynamic sub-fact of the form of a ship-bow rising up before our eyes. That is, using the identified sub-object of a ship-bow as a criterion, we are able to infer that the identification rule for the concrete object expressible by the singular term 'There is a ship there' is applicable to the object and that based on this the ascription rule for the predicate '…is approaching' applies to the property of approaching, which belongs

[8] To review the considered kind of inference, see Ch. II, sec. 9; Ch. III, sec. 10.

to the same whole object, both rules building the verifiability rule expressed by the statement

(b) That ship there is approaching.

Since this verifiability rule proves to be effectively applicable, it can be seen to have the higher-order property of being true, since a verifiability rule is the same as an e-thought-content. We see that based on the awareness of the dynamic sub-fact of a ship-bow rising we conclude that the dynamic grounding fact of the ship approaching is also real, that there is also a correspondence between the elements of the verifiability e-thought rule and those belonging to the grounding fact.

This is why we can still say that a thought expressed by *p*, its cognitive meaning, corresponds to the factual content *q*, even if it is a rule and its verifications are very often variable, partial and perspectivistic, relying on sub-facts. The sub-facts and their corresponding verifiability rules are like the branches of a tree that has a trunk, the grounding-fact, and its corresponding basal e-thought, its verifiability rule. Having access to some identifiable branches you can reach the tree. And a similar metaphor could be applied to the basal thought as a trunk and the derived e-thought-rules as the branches specularly corresponding to the first suitable external branches, the sub-facts, the last belonging to an external grounding fact, the external trunk. We will come back to this point in the end.

7. Formal definitions of truth

Assuming the suggested analysis of correspondence, we can symbolically express what could be called a formal definition of truth, giving us the logical structure by means of which we can identify the predicate '…is true' with the predicate '…corresponds with a fact.' As with the predication of existence, the predication of truth is of a higher-order. It is a *semantically metalinguistic* predicate applicable to thought-contents. We call a predicate semantically metalinguistic when it refers primarily to the *content* of the object language, contrasting it with a *syntactically metalinguistic* predicate, which refers only to the symbolic dimension of the object language. The statement

'"Themistocles won the battle of Salamis" is a historical statement'

can serve as an illustration. The semantic metapredicate '…is a historical statement' refers metalinguistically primarily to the semantic content of its

object-sentence, that is, to its thought-content-rule, and by means of this, secondarily, also to the real historical fact (acknowledging it as real). According to this view, *for any e-thought or content of belief* p*, to say that* p *is true is the same as to say that* p *adequates to an objectively real or actual factual content*. We can express this symbolically, using *p* to express the e-thought, replacing the predicate expression '…is true' with *T* and the predicative expression '... adequates to an actual fact' with *A*. The symbols *T* and *A* stand for semantic metapredicates belonging to a semantic metalanguage, by means of which they refer to the e-thought-content-rule expressed by *p*, which can be shown by placing *p* in quotation marks. Using '=' to express something like a (numerical) identity, here is a first identification of truth with adequation:

(1) *T* '*p*' = *A* '*p*'[9]

According to this identification, truth is the property of a thought-content expressed by a sentence *p* of adequating to some real-actual fact.

This formulation depends on the application of the monadic predicates '...is true' and '...adequates to a fact.' However, monadic predicates can often be unfolded into non-monadic predicates such as, for instance, '…is a father' into the more specific '…is the father of…' The same can be said of the predicates '…is true' and '…adequates to a fact,' which can be unfolded as more complete relational predications of a semantic metalanguage relating the thought expressed by *p* to the fact or factual content that *q* as '…is true for…' and '…corresponds to the fact that…' (*Cf.* Künne 2003: 74). We can also illustrate this point using an example. One could say

> '"Themistocles was the victor at the Battle of Salamis" expresses the same historical occurrence as "The Battle of Salamis was won by Themistocles",'

[9] We remember here Alfred Tarski's disquotational formula, according to which '"*p*" is true in L ≡ *p*.' Tarski's approach has the great merit of properly emphasizing the metalinguistic character of the truth-assignments in a formal language (*Cf.* Tarski 1944: 341-375). However, his formula does not overcome the philosophical problems of correspondence. If we replace sentence *p* with *Fa*, Tarski's theory does not provide criteria that tell us why we should apply *F* to the object referred to instead of to any other object, and it does not consider the necessity of criteria for the reference of the name *a*, which natural language requires. The task here is to review his insight in order to integrate it in our maximalist approach.

where '…expresses the same historical occurrence as…' is a relational semantic metapredicate primarily applied to the e-thought-content-rules expressed by the two object-sentences and only secondarily to the facts represented by them.

This means that the definition (1) can be more completely explained as stating that *for a given thought-content* p*, to say that* p *is true for the actual factual content* q *is the same as to say that the thought-content* p *adequates to the actual factual content* q. For this explanation, one can understand correspondence *as a relation of identity of contents expressed by* p *and* q, so that we can say that p = q. (I underscore *q* in order to show that its content, though also interpretable as an e-thought, is preferably interpretable as an actual or real fact in the world; how this is possible will be explained later…) To offer a simple observational example: suppose that the thought expressed by 'The Moon is white' is true. We only say this because of the real-actual fact that the Moon *is* white. And this is the same as saying that the e-thought-rule expressed by 'The Moon is white' corresponds to contents of observations of the white Moon understood as really factual.

Now, replacing the semantically metalinguistic relational predicate '…is true for the fact that...' for *T**, and replacing the also semantically metalinguistic relational predicate '...adequates to the fact that…' for *A**, we have the following more telling formalized version of a so-called formal definition of truth as adequation. In this definition, the e-thought-rule expressed by *p* and the actual factual content expressed by *q* are metalinguistically related by the metapredicates *T** and *A** as follows:

(2) '*p*'*T**'*q*' = '*p*'*A**'*q*'

More than an unpacking of (1), the identity (2) is a more complete formulation that individualizes the corresponding fact to be represented by any instance of *q*. According to (2), the assignment of truth is the same thing as the assignment of the relational property of correspondence, which can be viewed as the assignment of a qualitative *identity of content* between an e-thought-content-rule and an actual corresponding factual content. (As we saw, this identity of content should be analyzed in terms of structural isomorphism, added to the satisfaction of criteria for applying each component term of *p*…)

Finally, assuming that e-thoughts are verifiability rules, we can add that to say that an e-thought corresponds to a fact should be the same as saying that the verification procedure constitutive of the e-thought applies to a

fact. Symbolizing the semantic metapredicate '…is a verification procedure that applies to a fact' with *V*, we have:

(3) *T'p' = A'p' = V'p'*

More completely, symbolizing the dyadic semantic metapredicate 'given… the verifiability procedure effectively applies to the fact …' as *V**, we have:

(4) *'p'T*'q' = 'p'A*'q' = 'p'V*'q'*

These are, I believe, the best ways to represent in an abstract formal way the general identifications between attributions of truth, adequation and verifiability.

8. Negative truths

Now, consider a false singular predicative or relational statement *p.* Since it is false, such a statement does not correspond to any epistemologically objective real fact in the world. However, to say that *p* is false is the same as to refute the attribution of truth to *p*, which means to say that the statement *~p* is true. Here the problem arises. If *~p* is true and we accept adequation theory, it seems that *~p* must correspond to some fact. However, suppose that we replace *p* with the false statement (i) 'Theaetetus is flying.' In this case *~p* is to be instantiated by (ii) 'Theaetetus is not flying.' Then, at first glance it seems that we have in (ii) a true statement that does not correspond to any fact in the world! This would lead some to suspect that (ii) is true because it refers to a ghostly *negative fact*: the unworldly fact that Theaetetus isn't flying.

With the help of the preceding formulations, it is easy to reach a more plausible answer. That the statement that *~p* is true does not correspond to any actual fact in the world, even if it is instantiated by 'Theaetetus is not flying' and we know that Theaetetus is in fact sitting, since according to *~p* he could also be standing, lying down, running, etc. However, since *~p* means the same as '*p* is false,' and by saying that *p* is false one denies correspondence with an objective real-actual fact in the world, one denies that the verifiability rule has effective applicability in its proper context, and that is all. Despite this, as I have insisted, by imagining the false idea that Theaetetus is flying (that I symbolize as 'f'), we already accept that f corresponds with a possible fact, namely, with our imaginary dynamic fact of Theaetetus flying, which although epistemically objective in Searle's

sense isn't actually real in the sense of belonging to the external world. However, a possible fact can be no real external fact in any metaphysical sense; it is something that is located somewhere in the mind (-brain) when we imagine it. In summary, ~*p* and '*p* is false' only mean that *p* expresses a verifiability rule that although applicable to an only conceivable or imaginary state of affairs – a possible fact – does not effectively apply to any actual, objectively real fact.

Summarizing, if you consider a general statement like 'There is no cat with three heads,' it means the same thing as 'It is false that there is a cat with three heads.' And what this statement says is that although there is a corresponding conceivable factual-object that is a cat with three heads, there is no externally real fact-object in the world that is a cat with three heads. Still, one could argue that the statement that there is no cat with three heads is true because it agrees with the *fact* that there is indeed no cat in the world with three heads (Searle 1998: 393). However, here I must disagree. It seems more reasonable to think that this is a mere *façon de parler*, allowed by the flexibility of our natural language. The statement 'There is no cat with three heads' is true because it means 'It is false that there is a cat with three heads,' which says that there is no real fact constituted by a cat with three heads living somewhere in the outside world – only an imaginary one.

9. Self-referentiality

As expected, the identifications we have made until now also enable us to develop a kind of Tarskian answer to the so-called liar paradoxes of self-referentiality. Consider the following standard self-referential statement:

(i) This statement is false.

If this statement is true, what it states must be the case. But it states that it is itself false. Thus, if the statement is true, then it is false. On the opposite assumption, that the statement is false, then what it states is not the case, which means that the statement is true. Consequently, if the statement is true, it is false, and if it is false, it is true. This is the simplest example of a semantic paradox of self-referentiality involving the concept of truth, although there are many variations.

One of these variations is the indirect self-reference in which a statement refers to itself by means of another statement, generating the same paradox. Consider an example (Haack 1978: 135):

(1) The next statement is true… (2) The previous statement is false.

If statement (1) is true, then (2) is true; but if (2) is true, then (1) must be false... On the other hand, if statement (1) is false, then (2) must be false; but if (2) is false, then (1) must be true.

Having in mind our previous formal definitions of truth as correspondence, the general answer is that self-referential statements like these are mistakenly constructed because in all these cases the predicate '…is true' does not work as a semantically metalinguistic predicate referring to a complete thought-content. Rather, '…is true' functions *as a normal predicate built into the thought-content*, in this way belonging to the object language. Being mistakenly constructed, these statements have no proper cognitive meaning beyond their grammatical form. They might seem meaningful on the surface, suggesting that we should treat them as we would treat a statement with the form '*p* is true' or '*p* is false.' Once we have fallen into this trap, paradoxical consequences follow.

Now, why doesn't an affirmation like (ii) 'This sentence is true' generate a paradox? Consider the statement 'The sky is blue.' The truth-claim is here unnecessary, since implicit. For reasons of parsimony, a statement usually does not need the addition that it is true in order to be understood as expressing a truth. Because of this, the statement (i), though affirming its lack of effective applicability, naturally generates its truth-claim, since what it affirms (its falsity) is seen as though 'This statement is false' should be additionally true. The statement (ii), to the contrary, affirming its own effective applicability, though also devoid of content, resists a paradox-generating interpretation because the affirmation of its own applicability does not generate a statement that implicitly affirms its lack of applicability, adding to it its falsity.

Now, consider the sentence (iii) 'It is true that this sentence has nine words.' This is a perfectly normal true sentence referring to itself. Why? The reason is that the metapredication of truth is applied to the thought-content (verifiability rule) that the sentence in question has nine words *without* really belonging to this thought-content. For the same economical reason that assertions do not demand the explicit attribution of truth, (iii) is in fact understood as (iv) '"it is true that this sentence has nine words" is true', and this can be made more completely explicit as (iv) 'The thought expressed by the sentence "It is true that this sentence has nine words" is true.' This makes it clear that the relevant attribution of truth is not built into the relevant thought-content.

Furthermore, we can predicate the truth of a metalinguistic thought-content insofar as this semantic predication is meta-metalinguistic and so

on, since the e-thought, as an arrangement of apparently disembodied mental tropes, is also a fact.

10. Pragmatics of the correspondence relation

What we have seen up to this point was the frozen logical-conceptual structure of truth as correspondence. Now we will see how it works in the practice of truth-attributions, as a process occurring in time. The view I wish to defend here was inspired by Moritz Schlick's brief defense of the correspondence theory of truth (1910), though in my judgment this could be regarded as an empiricist revision of a relevant insight attributable to Edmund Husserl (*cf* sec. 31 of the present chapter). The idea is that correspondence has a pragmatic or dynamic dimension that deserves to be explored and cannot be captured in static formal definitions – an idea that should not sound strange to those who wish to combine correspondentialism with verificationism. We can begin by considering that very often we can establish an idealized sequence of (I choose) four successive moments, which we may call: (1) *suppositional*, (2) *evidential* (3) *confrontational* and (4) *judgmental* or *conclusive*. Together they constitute a very common form of verification procedure.

The best way to introduce the idea is by means of examples. Schlick used the example of Le Verrier's prediction of the planet Neptune's existence based on orbital perturbations of Saturn: Le Verrier first developed a hypothesis, which was later confirmed by observation, since the *contents* of both were the same. I next offer a more trivial example. Suppose that it is the rainy season in Northeastern Brazil, where I normally live, and that I ask myself: 'Will it rain in Natal tomorrow?' This is a suppositional moment. Now, when tomorrow comes, I open the door of my house and see that, in fact, it is raining heavily outside. This is the second, the evidential moment. Once I do this, I compare my earlier question with the observational evidence that it is in fact raining and see that the *content* of the question is like the *content* of my observation. This is the confrontational moment. Finally, considering that these contents are qualitatively identical (in fact, satisfying conditions (i) to (vi) of adequation), I conclude that the thought-content of my earlier hypothesis is true by adequation with the fact that today it is raining in Natal. This is the judgmental or conclusive moment. Now, if instead of seeing rain outside I see a very blue sky, the content of my observation contradicts that of my supposition. Seeing that the content of my observation in this proper context diverges from the content of the supposition, I conclude that p must be false: it is not raining in Natal today.

Examples like these are common, and an analogous procedure, as we will see, applies to non-perceptual truths. But for now, restricting myself to perceptual judgments, I can say that at least regarding cases like those considered above, we can formulate the following action-schema with four moments:

1) The *suppositional moment:* what I call 'supposition' can be a thesis, a hypothesis, a conjecture, a suspicion, a guess, a question, a doubt... In this first step we ask ourselves whether some thought-content-rule is true, that is, if the verifiability rule that constitutes it is not only imaginatively, but also definitely applicable in its proper context. We can express this as 'I suppose that *p*,' 'It is possible that *p*,' 'I guess that *p*,' 'Is it the case that *p*?,' where *p* expresses a content that can be perceived. This moment can be formalized as '*?p*' (call '*?*' the operator for supposition). This supposition is always made in the context of some linguistic practice.
2) What follows is the *evidential* or *perceptual moment*: the realization of a perceptual experience in an already more or less specified observational context gives us a perceptual content, which may or may not correspond to the content of the supposition.

Here we try to verify the truth of the supposition by finding a perceptual content that is identical to the content of the supposition. In the case of observational truths, this step is very simple. We look for an expected adequate perceptually reached content of thought that, in a suitable context, we simply read as a truth-maker (verifier), which can be rendered as 'I perceive the fact $\underline{o}$,' call it '*!*$\underline{o}$' (where '*!*' is the evidence operator). Phenomenologists have called this moment *registration* or *fulfillment* (*Cf.* Sokolowski 1974, Ch. 9). As we will see, there can be no question about the truth-value of *o*: it must be assumed as 'evidence' or 'certainty'. In fact, it must be *stipulated* as indisputable within the context of the practice, the language game in which it occurs; otherwise we would be daunted by the question of the truth of $\underline{o}$*!* which would also need to be grounded, leading us to an infinite regress. (The ontological problems concerning $\underline{o}$*!* will be discussed only at the end of this chapter.)

3) *Confrontational moment*: it is the comparison between the suppositional content and the factual content of the perceptual experience which makes possible the verification or falsification of the suppositional content.

Here we ask whether the supposition matches the evidential result of the perceptual experience. In the case I considered, I asked myself whether the thought-content-rule of the hypothesis was sufficiently similar to the factual content directly given to me in the perceptual experience (satisfying conditions (i) to (vi) of adequation). In the case of a perceptual experience, the positive answer can be summarized as $p = \underline{o}$. As will be better explained and justified later, here also we underscore o as $\underline{o}$, so that it can be read as either the thought-content-rule (a proposition) (o) or the actual factual content (presented by $\underline{o}$) fulfilling it, which involves an arrangement of external tropical criteria given in the contextually expected sensory experience. If the expected similarity of content between p and $\underline{o}$ is lacking, we have $p \neq \underline{o}$. (In its concrete details it is more complicated: as we already noted, usually the fact presented by $\underline{o}$ is only partially and aspectually experienced, which does not prevent me from saying, for example, that I *see* that it is raining all over Natal. Moreover, in practice it is often the case that we must have more than only one perceptual experience and in more than one way...)

4) *Judgmental* or *conclusive moment*: Finally, in the case in which $p = \underline{o}$, the thought expressed in the supposition will be accepted as *true*, otherwise it will be rejected as *false*. When $p = \underline{o}$, there is adequation and the conclusion is an affirmative judgment that can be symbolized as $\vdash p$. In the case in which $p \neq \underline{o}$, that is, in the absence of the expected adequation, the thought p is false. This can be expressed by the negative judgment symbolized as $\vdash \sim p$.

Now we can summarize the four steps of this whole verifiability process regarding the discovery of perceptual truths of the simple kind considered above in the following temporal sequence:

$$?p,\ !\underline{o},\ p = \underline{o} \ /\ \vdash p$$

This analysis shows that in many cases one finds adequation (particularly as identity of content) between some suppositional e-thought-content-rule $?p$ (which is only a considered or imagined verifiability rule in its possible application) and some perceptual e-thought-content-rule $!\underline{o}$ (given by the definitely applied verifiability rule) that within the linguistic practice in which it is stipulated is regarded as indisputable. In other cases, the adequation is only between the supposition and an imagined, non-actualized fact, being therefore distinct from what can be found in the observation. In these cases, the statement must be false.

It is also worth noting that the standard statement of ⊢*p* (a judgment) has the form of the *report of an assertion* that is settled. However, this assertion can always be questioned again. In this case, new verifying procedures can reconfirm the judgment or detect some inadequacy refuting it in an at least virtually interpersonal way (*Cf.* Sokolowski 1974, Ch. 9).[10]

Now, how can we understand the adequation relation as a qualitative identity of content (structural isomorphism, identity of cognitive rules, intentionality…) in terms of the application of verifiability rules? Here is my suggestion. When I first perceive that it is raining in Natal, the indexical phrase 'now in Natal' expresses the building and application of an indexical identification rule of a spatiotemporal region to which the predicate '…is raining' is applied. This predicate expresses an ascription rule definitely applicable to the region by the satisfaction of configurations of tropes constituted by the countless drops of water falling from the sky above. This combination of satisfactions gives me the arrangement that constitutes the sub-fact that is the truthmaker which allows me to infer the content building the grounding fact o! that it is raining in (all parts of) Natal today. Now, p = o means that the contents of both e-thought-rules are identical. In more detail, there is an adequation between both e-thought-content-rules or, in still more detail, the identification rule of *p* has a one-to-one relation with the identification rule of o, the ascription rule of *p* has a one-to-one relation with the ascription rule of o, the concatenation between the rules of *p* and of o is the same, there is categorical match, intentionality and causality; p is intended to fit o, and o has a causal direction of fit concerning *p*, since it makes *p* true. Consequently, the verifiability e-thought-content rule *p* adequates to the verifiability e-thought-content rule *o*, even if in details this can occurs by means of the most diverse sub-factual isomorphic matches of criterial configurations.

Now one could object: must we have a *qualitative* identity between *p* and *o*? It is true that between the ?p of yesterday and even the ?p that I made to myself as I awakened today and the !o there is indeed qualitative identity. However, I cannot believe that at the moment when I perceive that it is raining, p and o are qualitatively distinct. It seems to me more plausible that the identity p = o in the perceptual moment have a numerical identity, which means that Husserl was in his own way right in

[10] An at least virtual interpersonal confirmation is here important. In my view, truth must be able to ultimately satisfy an interpersonal consensus made authentic by its achievement through adequate agreement within a *critical community of ideas* (a community with equally competent members, with the same rights of interaction, etc.), a point particularly relevant in regard to the collective acceptance of complex law-like generalizations (*Cf.* Habermas 1983).

understanding correspondence as a form of identity (See sec. 31 of this chapter). Moreover, it is always possible to interpret $\underline{o}$ as a real external fact and not propositionally, as we can do with the mere identification $p = o$.

11. Retrograde procedures

Now, what was presented above is what we may call an *anterograde* way to achieve truth. I call it so because we went in a temporal sequence from the supposition containing a conceivable e-thought-content-rule to the perceptual evidence that confirms the supposition by the application of a perceptual e-thought-content-rule that is qualitatively identical with the supposition. However, a move in the opposite direction is equally feasible. We can have a truth-value attribution that has its origin in perceptual experience, progressing from evidence to the affirmation of a supposition – a way to discover truth that I call *retrograde*.[11]

Here is a simple example of a retrograde verification procedure. I open the door of my home in Natal with the intention of going out and unexpectedly see that it is raining. Since I need to go out, I go back inside to look for an umbrella, aware that it is raining… In this case, the perceptual evidence comes first. However, it seems clear that the recognition of truth does not occur as a direct product of sensory experience since I could see rain without consciously perceiving it. This suggests that the initial rough and pre-conscious sensory-perceptual state was different from the state of awareness that immediately followed, namely, the conscious awareness that it is raining. (Suppose I open the door to get some fresh air although I see I do not even pay attention to the fact that it is raining outside. If someone then asks me if it is raining, I will pay attention to the already non-reflexively roughly applied conceptual rule for rain, compare it with a similar now fully conscious application of this rule and answer in the affirmative). Thus, it seems that we can explain the process of arriving at the truth included in the judgment of the given example in the following way: First, I have the rough, non-reflected observational experience '$\underline{o}$!' of rain. This momentary experience makes me immediately recall the fully conceptualized ascription rule for 'it is

[11] I believe the anterograde and retrograde procedures are a more explicit version of a distinction already present in Husserlian phenomenology: the distinction between 'truth as correctness' (*Wahrheit als Richtigkeit*) and 'truth as discoveredness' (*Wahrheit als Entdecktheit*) respectively (See Sokolowski 2000, Ch. 11).

raining,' which together with the identification rule for 'the city of Natal today' forms the supposition '?p.' Finally, I compare the content of my first observation with the content of this recalled e-thought-rule of raining in Natal today. Once I see that o = p, I am led to the conclusion that it is true that it is raining or ⊢p. If I am right, then this process is normally completed very quickly, which accounts for our lack of awareness of its different steps. Anyway, this is a *retrograde* discovery of truth, which I believe that can be summarized in the following sequential formulation:

!o, ?p, o = p / ⊢ p

Clearer cases of retrograde awareness of truth occur when we have an unexpected sensation or perception that we only slowly come to be aware of in its true conceptualized nature. To illustrate I give two examples. The first is related by Paul Feyerabend in his autobiography. He writes that once when he was sleeping he at first mistakenly identified a feeling with a cramp, and only when he woke up did he see what he was really feeling: a severe pain in his leg. We may call the first sensation '!s,' mistakenly taken as a cramp or s'. In the process of waking up, he must have been led to recall the most appropriate conceptual rule for pain as '?p.' As he clearly identified s with p, he realized that he was feeling pain in his leg, reaching the conclusion ⊢p.

The second example is of an experience that I myself once had. A nice woman gave me a teacup at her home containing a sweet beverage, without saying what it was. I was sure I knew the taste, though I could not identify it. Hence, I must have applied a mugh ascription rule, which I call !t. However, since the context gave me no clue as to what the liquid in the cup was, I needed about a minute to recall the taste of juice from pressed sugarcane, that is, '?p.' Then, by comparing this conceptual memory ?p with the taste of the liquid !t in the cup and seeing that t = p, I came to the obvious conclusion: the liquid was pressed sugar-cane juice. Here the action-schema is:

!t, ?p, t = p / ⊢p.

The retrograde procedure seems to be the inverse of the anterograde, also because the first moment of both seems loose, unsettled, insufficiently determined.

12. A more complex case

The cases I have considered until now are the simplest sensory-perceptual ones. However, the pragmatics of adequation can be extended to the truth of non-observational e-thought-rules, which I will here call *mediated thought-contents*. Suppose that Lucy is at Charles de Gaulle Airport in Paris, waiting to board a flight to Dakar. The flight lasts approximately five hours. She calls her daughter, who lives on a farm in Senegal and asks her how the weather is in the city of Dakar. She wonders if it is sunny. This is supposition ?p. Suppose that after a while her daughter answers that the weather in Dakar is and will remain mild and warm enough. There is no significant reason for doubting this information, which she takes as providing adequate evidence. The factual thought-content expressed by '!q' that she had after she heard about the weather in Dakar is the same as the thought-content belonging to her hopeful question '?p.' Consequently, since p = q, she concludes that p is true, that the weather in Dakar is and will remain mild. But the thought-content-rule expressed by !q is not observational! It is the result of testimonial inferences that are unknown to Lucy. Suppose that her daughter got this information from her husband, who had read a weather report, and that this information had its origin in meteorological observations of weather conditions around Dakar. In this case, putting '>>>' in the place of some chain of reasoning unknown to Lucy that led to the factual judgment expressed with !q, and putting '!o' in the place of the observational meteorological thought-contents that in some way led to !q (which will probably be similar to those that she will have when she arrives in Dakar five hours later), we can formally structure the verification process in which p is presently made true for Lucy as follows:

?p, (!o >>> !q), !q, p = q / ⊢ p

Important to note is that the evidential character of the observation !o is taken as preserved in the supposed inferential chain that leads to !q (I put the process in parentheses in order to show that it is unknown to Lucy and even to her daughter). The informational content is transmitted from thought-content to thought-content up to the conclusion !q, which inherits the evidential character of !o, and then !q is compared with the question expressed by ?p. Thus, contrary to our most natural expectation of how adequation should work, the truth of ?p isn't directly confirmed by the observational fact represented by !o, *but by something derived from it, namely, by* !q, understood as also representing a fact, a personally non-

experienced state of affairs in the world. The adequation is between unfulfilled and fulfilled thought-content rules, the last ones also understood as being fulfilled by a factual content composed of external tropical arrangements.

The foregoing example is one of an anterograde verifiability procedure, beginning with one supposition (the question) and ending with a comparison between the supposition and a *derived* evidential thought-content of an unexperienced fact. However, we may also have a retrograde procedure with a chain of reasons that ends up by matching a derived piece of evidence with a supposition. So, imagine that at the beginning of the flight to Dakar the pilot informs the passengers that the weather in Dakar will be mild and warm enough. Each passenger will be led to the conclusion that the weather in Dakar will, in fact, be mild by means of another indirect and for them also unknown evidential chain. However, in this case, it is the evidence that recalls the concern regarding weather conditions. This concern is satisfied by means of a comparison of contents from which the final judgment results that the weather in Dakar will be mild. This retrograde process can be summarized in the following temporal sequence:

$$(!\underline{o} >>> !\underline{q}),\ !\underline{q},\ ?p,\ \underline{q} = p \mathrel{/\!\vdash} p$$

We see that the opposition between anterograde and retrograde verification repeats on mediated levels. We may guess whether the intuitions of some researcher who still does not know how to prove some hypothesis, though having a glimpse of its truth, depends on unconsciously noticing that the knowledge of some factual content expressed by $!\underline{q}$ might be derived from evidential observations or postulates.

13. General statements

General statements of e-thought-contents – universal and existential – are also involved in the pragmatic process of adequation, as an identity between the contents of the hypotheses and contents of sets formed by the respective conjunctions and disjunctions, often resulting from inductive inferences ultimately based on observational facts. So, suppose that $\vdash p$ is the assertion: 'All the books on this shelf are in English.' Further, suppose that I reach this generalization casually in a retrograde form from earlier observations '$!\underline{o}_1$, $!\underline{o}_2$… $!\underline{o}n$,' of each book on the shelf. The action-schema is the following:

$$\{!\underline{o}_1 \,\&\, !\underline{o}_2 \,\&\ldots\, \&\, !\underline{o}n\} \rightarrow !\underline{q},\ ?p,\ \underline{q} = p \,/\!\vdash p$$

Of course, it can be different. It can be that I first ask myself if all the books on the shelf are in English. Then I look at each of them, concluding in an anterograde procedure that this hypothesis is true:

$$?p,\ \{!\underline{o}_1 \,\&\, !\underline{o}_2 \,\&\ldots\, \&\, !\underline{o}n\} \rightarrow !\underline{q},\ p = \underline{q} \,/\!\vdash p$$

Now, suppose that for another Mrs. Hildish asks: 'Is there at least one book in Italian on my shelf?' Now, after searching, she finds just one. We call it '$!\underline{o}_1$.' This enables her to affirm that there is at least one book in Italian on her shelf, concluding by means of an anterograde procedure:

$$?p,\ \{!\underline{o}_1 \vee !{\sim}\underline{o}_2 \vee\ldots\vee !{\sim}\underline{o}n\} \rightarrow !\underline{q},\ p = \underline{q} \,/\!\vdash p$$

As in the previous cases, this example deals with a general deductive conclusion, but it is easy to see that inductive generalizations should also have similar structures, given that they are also restricted to some more or less vague domain (See Appendix to Chapter V, sec. 3).

Now we return to the old question of knowing if there must be general facts – the *all facts* – over and above singular facts (Russell 1918; Armstrong 1997, Ch. 13; 2004, Ch. VI). Bertrand Russell, who seems to have discovered the problem, defended their existence as follows:

> I think that when you have enumerated all the atomic facts in the world, it is a further fact about the world that *those are all* the atomic facts there are about the world, and that is just as much an objective fact about the world as any of them are. It is clear, I think, that you must admit general facts as distinct from and over and above particular facts (Russell 1956: 236, my italics).

It seems to me that this is much more a worldly question than Russell supposed, since it can be shown that his *all fact* is not a fact hanging over any other. In the examples above, all that is needed to get the totality of facts is an additional *limiting fact* restricting the extension of the generality, first to books belonging to my first shelf and then to books belonging to Mrs. Hildish shelf. I agree that descriptions of such limiting facts need to be added to the given sequences of particular conjunctions or disjunctions in order to close their domain. But these limiting facts are nothing but ordinary empirical ones. And the harmless affirmation 'those are all', meaning 'there is nothing beyond these' can be inferred as a consequence of adding the conjunction or disjunction of the singular facts

to the corresponding empirical singular limiting facts, in the given case the facts established by the spaces the shelves have for their books! Using a still simpler example, if I say that I have only three coins in my pocket, the 'all fact' is given by the domain established by the fact that there is a pocket in my pants that I use to carry coins. Moreover, the only difference between the examples given above and an extensive fact like 'All men are mortal' is that the delimitation of the last domain is probably the whole earth during the whole existence of the species *Homo sapiens*, which is a much larger and more vaguely delimited domain. This is how Russell's mysterious and inconvenient *all fact* disappears.

14. Some funny facts

There are a variety of puzzling 'funny' facts, and I will only select a few to give some indicative explanations. One of these is that of self-psychic (self-reported) truths. It is easy to know the truth-value of the thought *p*: 'I am in pain.' I believe that here as well there is adequation. But first, I need to learn the rule. A first step to this is that I interpersonally learn to identify the location of pain. Then, helped by a considerable network of other concomitant, previously and later observable occurrences, along with the fact that I am told by others that pain is none of these, I discover, by means of *induction by exclusion*, that pain must be an invisible but physically located feeling of intense discomfort… Even if others cannot have interpersonal access to the subjective feeling of my pain in order to confirm it, I am able to make my verifiability rule for pain highly plausible, even if the logical possibility of interpersonal access to my pain itself cannot be excluded.[12] Now, suppose that I have a headache. The first thing I have is an unnamed feeling of pain: '!s.' Then comes '?p': the actualization of the memory of what the feeling of having a headache means (the conceptual rule), which is what I associate with the word. Then I make the identification s = p, being led to the conclusion ⊢p:

!s, ?p, s = p / ⊢ p

Here I discovered the truth that I have a headache in a retrograde way. An anterograde way to reach the same truth would be the case of a woman who guesses that she will have a headache because she has drunk red wine, and she knows she always has a headache after drinking red wine.

[12] See my objections to the private language argument in Chapter III, sec. 13 of the present book.

Wittgenstein offered, as is well known, an expressivist explanation for such cases. For him the utterance 'I am in pain' is nothing more than an extension of natural *expressions* of pain like 'Ouch!' (Wittgenstein 1984c, I, sec. 244). In this case, our schema would be something like '!s ├ *p*' without adequation. I do not reject this possibility. But I find it easier to believe that this could be the expression of a more direct reaction that turns out to be seen as true only *after* the exercise of the previous, more elaborate cognitive process of induction by exclusion concerning auto-psychic states and induction by analogy concerning the hetero-psychic states (Costa 2011, Ch. 4).

Another odd case is that of true counterfactual conditionals. Consider the statement (i) 'If Evelyn were the queen of England, she would be a public figure.' The objection is that there appears to be no fact that can make this sentence true, since Evelyn isn't the queen of England. However, statement (i) seems to be true! Nevertheless, the solution is easy. Although there is no actual fact that can make statement (i) true, this is not what the conditional requires. What statement (i) requires as its verifier is not an actual fact, but only a *possible fact*. The possible or conceivable fact that makes the statement true is that under the *assumption* that the antecedent were true, namely, that Evelyn is in fact the queen of England, the truth of the consequent will be unavoidable, that is, she will surely be a public figure. That is, the truthmaker of (i) is a *modal fact* that could also be expressed using the vocabulary of possible worlds. In other words, suppose that *We* is any nearby possible world where Evelyn is in fact the queen of England. Since in our world all queens of England are public figures, we can infer that if someone is the queen of England in *We*, this person will also certainly be a public figure. Assuming that Evelyn is the queen of England in *We*, she is also (certainly) a public figure in *We*. We conclude that it is certainly true that if Evelyn were the queen of England she would (certainly) be a public figure, because the expressed thought-content certainly *corresponds with the expected fact* belonging to a conceived counterfactual circumstance given in *We*. Understanding (i) as the supposition ?p, and calling the certainty that in any nearby possible world Evelyn would be the queen and therefore a public figure q, we can summarize the anterograde process as follows:

?p, (*We*)q, p = q, / ├p

A second similar example is, (ii) 'The Dalai Lama never slept with a woman, but he could have.' This is certainly true because it means the same thing as (iii) 'Although the Dalai Lama never slept with a woman in

the actual world, there is a nearby possible world *Wd* (our world with some differences) where he slept with a woman.' The statement (iii) is true, since it corresponds to the conjunction of an actual and a possible (conceivable) fact, this conjunctive fact being conceivable as a highly probable physical possibility (ontologically, an association of actual and possible tropical arrangements).

One could also ask about ethical truths. Consider the statement (iv) 'Dennis should help the drowning child.' Suppose that despite being a very good swimmer, Dennis didn't even try to help the drowning child, because he is a sadist. We would not be inclined to say that (iv) is true, but rather that (iv) is *right*. It is right in a similar way as an illocutionary act like 'I promise to go to your anniversary celebration' can be felicitous. The statement about Dennis would be morally right because it is in conformity with a utilitarian norm, let us say, the rule according to which:

> UR: One should help another person in mortal danger, insofar as one does not put oneself in real danger.

Note that what counts in this case is not truth, but *normative correctness* – adequation with a norm, though the mechanism of validation is similar. Statement (iv) is validated by what could be called the moral *norm* UR (an equivalent to the *fact* regarding truth). Finally, there is still the case of the validity of such utilitarian norms. In an attempt to achieve this, consider the following utilitarian normative principle:

> UP: A morally correct rule is one that when applied under normal circumstances brings the greatest possible amount of happiness to all participants, without significant unhappiness to anyone.[13]

Suppose it is a *fact* that when people act in accordance with this principle the well-being of their whole community increases. Assuming that this is our ultimate goal, this principle can be considered correct or true, and we can say that UP validates UR, which validates (iv). (Note that the normative principle UP as much as the norm UR are moral facts that should be also instantiated as arrangements of tropes.)

---

[13] My preferred moral theory is two-tiered utilitarianism. According to this view, we should apply rule-utilitarianism in ordinary situations, although in extreme situations, utilitarian rules are defeated and we must turn to act-utilitarianism. (Hare 1981, Ch. 2)

Obviously, this is just an illustration. The greatest problem faced by ethical statements is the same as with any other philosophical statement. Unlike the statements of natural sciences, they belong to those speculative domains wherein we are only able to make the truth of our statements *more or less plausible*.

15. Expansion to formal sciences

Analogous logical structures and dynamic procedures can be found in the formal sciences, allowing us to generalize adequation theory to a domain traditionally claimed by coherence theories of truth. The main difference is that while for empirical truths inferences are mainly inductive, for formal truths they are normally understood as deductive. Suppose we want to demonstrate that the sum of the angles of any Euclidean triangle is 180°. We can do this by first proposing that this could be the case: '*?p*' and then searching for proof. One proof would proceed by drawing a straight line through one of the vertices of the triangle, so that this line is parallel to the side opposite to this vertex. Since the three juxtaposed angles formed by the parallel and the vertex of the triangle are the same as the internal angles of the two opposed vertices of the triangle plus the angle of the first vertex, and their sum is obviously 180°, we conclude that the sum of the internal angles of this and indeed of any Euclidean triangle must be 180°. This deductive conclusion is the evidence '*!q*' – the truthmaker as a *geometrical fact* constituted, I suppose, by geometrical tropes (*Cf.* Appendix of Chapter III, sec. 4). Since we see that the content of *!q* is the same as the content of the hypothesis *?p*, we conclude $\vdash p$. Using '*as*' for the axioms or assumptions (the formal data), the form of this anterograde procedure can be summarized as:

$$?p,\ !as >>>\ !q,\ p = q,\ / \vdash p$$

It is important to see that *!q*, stating the fact that makes the thought-content *p* true, as in the case of Lucy's question, should not be placed at the beginning, but at the *end* of a chain of reasoning. Unlike Lucy, a geometrician can (and should) go through the whole procedure.

Now, an example from mathematics: we can prove the arithmetical identity statement (i) '2 + 2 = 4' in a Leibnizian manner.[14] We begin with definitions (which here correspond to basic perceptual experiences in empirical sciences). First, we define 2 as 1 + 1, 3 as 2 + 1 and 4 as 3 + 1.

[14] Leibniz' original proof can be found in his 1765, liv. IV, Ch. 7, Sec. 10.

We call this set of definitions '*d.*' Replacing in statement (i) the numbers 2 and 4 with their *definiens*, we get (ii) '(1 + 1) + (1 + 1) = (3 + 1).' Since 3 is defined as 2 + 1, and 2 as 1 + 1, we see that 3 can be replaced by (1 + 1) + 1. Now, replacing the number 3 in its analyzed formulation in (ii), we get the *arithmetical fact* represented by (iii) '(1 + 1) + (1 + 1) = (((1 + 1) + 1) + 1),' which is the same e-thought-content as '2 + 2 = 4.' In this way, we have derived confirmatory evidence for the hypothesis '*?p*' posed by statement (i), which is the (supposedly tropical) factual content of '*!q*' described in (iii). This confirmatory evidence serves to check the hypothesis '*?p*' that 2 + 2 = 4. Again, abbreviating the definitions as '*d*,' we have the following anterograde verificational action-schema:

$$?p,\ !d >>>\ !q,\ p = q\ /\!\vdash p$$

Once more we see that the factual content expressed by the identity *!q*, which serves to check the hypothesis *?p* that 2 + 2 = 4, is not the same as the definitions of 1, 2, 3 or 4, as might be initially assumed. It is the *result* of a deductive reasoning process based on these definitions, a reasoning process deductively derived from its definitional premises. This result, expressed by *!q,* represents the arithmetical fact represented by the supposition *?p*, so that $p = q$, which makes p true.

Finally, we can give examples involving logic. Consider the following theorem of modal logic: $P \rightarrow \Diamond P$. This can be seen as our hypothesis *?p*. How do we prove it? In the S5 modal system, we can do this by using as assumptions the axioms AS1, $\Diamond P \leftrightarrow \sim\Box\sim P$, and AS3, $\Box\sim P \rightarrow \sim P$. Taking these axioms and a few rules of propositional logic as the evidence '*as*' we construct the following anterograde proof of the theorem:

The hypothesis is: '*?p*,' where $p = P \rightarrow \Diamond P$

The proof:

1	$\Box\sim P \rightarrow \sim P$	(AS3)
2	$\sim\sim P \rightarrow \sim\Box\sim P$	(1TRANS)
3	$P \rightarrow \sim\Box\sim P$	(2~E)
4	$\Diamond P \leftrightarrow \sim\Box\sim P$	(AS1)
5	$\sim\Box\sim P \rightarrow \Diamond P$	(4 ↔E)
6	$P \rightarrow \Diamond P$	(3,5 SD)

Now, the conclusion (6), $P \rightarrow \Diamond P$, is the '*!q*,' which represents the derived logical fact that serves as a verifier for *?p*, and since $p = q$, we conclude

that *p* is true, that is, ⊢ *p*. Using our abbreviation, we get the following anterograde verificational action-schema:

?p, !as >>> !q, p = q, /⊢ p

Since the *logical fact* represented by *!q*, which carries with it evidence derived from the assumed axioms, is presented by the same e-thought-content-rule as the hypothesis *?p*, we conclude that we have adequation. We conclude that *p* is true, or ⊢ *p*. – Also relevant is to note that in the case of formal facts we do not need to underline statement letters like *a* or *d* or *q*: there is no need to distinguish between the conceived and the real-actual facts, since here both can be regarded as the same.

Of course, one could also find a retrograde form regarding any of the three above exemplified cases. Considering only the first, suppose that someone, having the strong intuition that the sum of the internal angles of an Euclidean triangle is 180°, decides to draw a straight line that touches the vertex of a triangle, this line being parallel to the opposite side. This person could then easily prove that this triangle and in fact any Euclidean triangle would have 180° as the sum of its internal angles. But in this case, the person would have the following retrograde verification procedure:

!q, !as >>> !q, ?p, q = p, /⊢ p

The *!q* would work here as the insight into the truth of a conjecture, something to be compared with an unexpected observation.

The upshot is that the procedures with which we demonstrate the adequation of formal truths are structurally analogous to the procedures with which we demonstrate the adequation of empirical truths. Even so, there are some differences. The most obvious is that formal truths are deductively inferred, while empirical truths unavoidably include inductive inferences.

16. Why can analytic truths be called true?

Finally, we can apply a similar procedure to analytic-conceptual statements, showing that they are also called true because of adequation, even if this is a limiting-case. It is possible to say, for instance, that the analytic statements ‘It is raining or it is not raining’ and ‘Bachelors are not married’ are true because they correspond to the respective facts that assuming the principle of the third excluded it must be either raining or

not, and that by definition it isn't possible for a bachelor not to be unmarried. But to what extent are we entitled to say this?

Assume first, as we did in our objections to Quine's argument against analyticity, that analytic statements are true due to the proper combination of the component senses of their expressions. In this case, our question is: are there facts that make analytic statements true? And if they exist, how do they make these statements true? To find an answer, consider the following analytic statements:

(1) Either it is raining or it is not raining.
(2) If John is the brother of Mary, then Mary is the sister of John.
(3) Bachelors are males.
(4) A triangle has three sides.
(5) A material body must have some extension.

Surely, these statements are all true in themselves: if there is a fact making them true, it is not a fact in the world. However, we are still allowed to say that they are made true by logico-conceptual, conventionalized *facts*. Thus, statement (1) is made true by the logical fact that '$\varphi \vee \sim\varphi$' (the law of the excluded middle), which it instantiates. Statement (2) is made true by the conceptual fact that the brother-sister relation is reflexive. Statement (3) is made true by the conceptual fact that a bachelor is conventionally defined as an unmarried adult male. Statement (4) is made true in Euclidean geometry by the conceptual fact that a triangle can be defined as a closed plane figure with three straight-line sides. And statement (5) is made true by the conceptual fact that part of the definition of a material body must include the requirement of some spatial extension. These are conceptual facts supposedly instantiated by arrangements of our mental tropes and their combinations.

In all these cases the statements are self-verifiable, that is, the intertwining of rules that constitutes the verifiability rule of an analytic statement is verified not by its application to the world, but by means of an application of one rule to the result of the application of the other in a way that makes the whole true independently of any state of the world. For instance, $\sim$(P & $\sim$P) is tautologically verified by its truth-table, in which we combine the rules for the application of the negation and the conjunction in ways that always gives as a result the value true.

Moreover, we can summarize this process of self-verification of the above statements by applying the same action-schemata we did with the statements considered in the last section. Thus, in case (1) we can begin with the question $?p_1$ = 'is it the case that it is raining or not raining?'

Faced with this, we immediately realize that the sentence instantiates the principle of the excluded middle or '$\varphi \vee \sim\varphi$', and that this instantiation, like any other, can be symbolized as the instantiation of the logical truth or fact represented by '$!p_2$,' which is proved true by the application of a truth-table to the sentence. This suffices to make $?p_1$ true, because we can see that independently of any sense given to its constituent parts, its logical structure warrants its truth. We can summarize the self-verifying action in which we find the adequation in the same anterograde way as in the first of our examples:

$$?p_1, !p_2, p_1 = p_2 / \vdash p$$

Putting differently: in this case, the thought-content is identical with an instantiation of a logical truth of propositional logic, a logical fact that makes (1) true by self-verification.

In other cases, reasoning may be necessary. In case (3) the suppositional moment '$?p_1$' is: 'Are all bachelors males?' To verify this, we first need to take the definition of a bachelor as our point of departure: '!d' (*Df.*) = 'A bachelor is an unmarried adult male.' From !d we can infer the conceptual fact $!p_2$ = 'All bachelors are males.' Summarizing the steps of this anterograde self-verificational procedure, we get:

$$?p_1, !d \rightarrow !p_2, p_1 = p_2 / \vdash p_1$$

It is correct to say that analytical thought-contents are true by courtesy, since they cannot be false. But despite this, it is not senseless to speak of their truth as correspondence or adequation with facts. The reason is clearer in cases like the last one. For even if these cases are all ones of self-verification, the procedure is not always direct and transparent, often requiring a reasoning process.

Finally, what about contradictions like (6) 'It is raining *and* it is not raining'? Suppose we call the statement of this contradiction the supposition '?p,' which is shown to be opposed to the true statement '~p,' asserting the factual content that it cannot be the case that it is raining and simultaneously not raining at the same time and place, which is derived from the principle '!*q*' of non-contradiction: $\sim(\varphi \,\&\, \sim\varphi)$. In this simple case, the anterograde verifying procedure will be:

$$?p, !q, q \rightarrow \sim p, p \neq \sim p, \vdash \sim p$$

The conclusion is that p is false, since the principle of non-contradiction shows that p cannot be the case and that strictly speaking there can be no fact in the world able to verify p. The verifying procedure that falsifies p is the self-falsifying cognitive action that gives the contradiction its contradictory meaning.

17. The insufficiency of coherence

That truth has something to do with coherence is beyond doubt. If Mary says that she was breathing while she was asleep last night, we accept her statement as obviously true. We believe Mary, even if we did not watch her sleeping last night, because her statement is coherent with our accepted belief-system. We are certain that people will die within a few minutes if they cannot continuously breathe oxygen. If Mary tells us that she visited the Moon while asleep last night, almost everyone would consider this statement to be false, because it clashes with the generally accepted commonsense understanding of what is possible or impossible under ordinary life circumstances, together with our system of scientifically confirmed beliefs. Coherence is obviously related to truth, and according to most coherence theorists, a belief is truer the more it is integrated into our system of beliefs, which also means that truth is a question of *degree* (e.g., Blanshard, 1939, Ch. XXVII).

Bernard Bosanquet (2015: 24) once gave an interesting example intended to show that a greater amount of supporting information makes a statement *more true*, which seems to vindicate the idea that some kind of integration of a statement within a system of beliefs is what makes it true. He noted that the sentence 'Charles I died on the scaffold' seems quite true when said by a leading historian and far less true when said by a mere schoolboy. The child has at most a name and a picture in his mind, while the historian knows from documents and historical studies a wealth of meanings associated with the sentence (*Cf.* also Blanchard 1939, Ch. XXVII, sec. 4-5). The aim of this example is to show that increasing the coherence of a statement increases its degree of truth.

Nevertheless, there is an alternative interpretation. We can say that the example only shows that the historian's claim to know the truth has a better chance to be confirmed. In other words, it is his *truth-holding* (*Fürwahrhalten*) *that has a higher chance of achieve truth*. This alternative is better, since there is no indication that our ordinary view of truth has degrees. Hence, the example only confuses the degrees of probability that a person knows the truth – the probability of truth-holding

that can be attributed to the person – with an illusory degree of truth in itself.

The best known objection to the coherence theory of truth is the following. Since countless possible belief-systems can be constructed, any proposition *p* could be true in one system and false in another, violating the non-contradiction principle. This objection, however, was never regarded as a major difficulty by coherence theorists (e.g. Bradley 1914; Blanshard 1939, vol. 2: 276 f.; Walker 1989: 25-40).

One could, for instance, answer the objection that some thought-content *p* can be true in one system and false in another in a way that eliminates the contradiction. One can introduce the idea of the *system of all systems*, namely, the most encompassing system of beliefs agreed upon by a community of ideas at time *t* (preferably the best informed and trained community that we are able to consider…). To this can be added the fundamental subsystem contained in the system of all systems, which is *the real-world belief-system*, so that this system generates all the other derived sub-systems that might fall under the epithet 'fictional.' The novel *Madame Bovary*, for instance, is for us a fictional subsystem belonging to the all-encompassing system of systems. That at the end of the novel Charles says, '*C'est la faute de la fatalité*,' is true *in* the context of the novel, but false for the real-world system, because in our real world there was never any Charles Bovary married to Emma Bovary and able to say this sentence regarding the series of events that led up to her suicide. The admission that Charles made this comment is thus true in the novel and false in the real world, which does not lead to a contradiction, not only because these are two belief-systems, but also because they do not conflict, as what counts is the real-world system, where this sentence was never uttered in a proper context.

Consider now a second example, the statement that the sum of the angles of a triangle is 180°. This is true in the system of Euclidean geometry, but false in Lobachevsky's and Riemann's systems. And it is in the end false regarding the physical real-world system. Consider, finally, the statement that the value of a good is determined by the importance people assign to it as a means to achieve their desired ends. It is considered true in the subjective economic theory of value and false in the labor theory of value, since for the latter the value of a good is determined by the amount of labor required to produce the good... Nonetheless, regarding the real-world system, the first theory seems to be (according to the great majority of economists) more probably true.

Surely, this view relativizes truth to a certain extent, by limiting it to a time and a community of ideas, making truth-theory to a certain extent

subordinate to our *taking things to be true* (*das Fürwahrhalten*).[15] However, in the end this would not be a problem if we agree that 'the truth,' that is, *absolute* truth, is actually nothing but a kind of *directive idea* that helps us evaluate our holding something to be true, but has no decisive identity with what we normally accept as true or false. – As already noted (Ch. IV, sec. 30), even if by chance we were to discover an absolute truth, we would not be able to know with any certainty that we had really found it (See Popper 1972, Ch. 2). That is, when we say that *p* is true, we only *assume* that *p* is the final truth until we find some sufficiently good reason to falsify *p* (if *p* is empirical) or abandon *p* (if *p* is a formal statement). Because of this, a true theory of truth is a theory of what leads us to *take* something to be true rather than a theory of absolute truth. The same can be said regarding the concept of knowledge. We pragmatically treat our truths and knowledge of truths as if they were the ultimate ones, simply postulating or assuming we have achieved final truths and knowledge. But concepts like those of final, ultimate or absolute truth and knowledge can serve only as *directive ideas*. They are 'as if' concepts since they cannot possibly have experienceable objects that allow us to see if we have achieved them.[16]

A strategy like that of admitting a system of all systems that includes a real-world system as the most fundamental seems to overcome the objection of contradiction. Nonetheless, even so coherence theory remains problematic, since the insurmountable problem of this view is located elsewhere. I call it *the problem of circularity*.

The problem of circularity arises when we try to define coherence. Traditionally coherence has been conflated with *consistency*. A set of propositions (thought-contents) is said to be consistent when the

---

[15] I say 'to a certain extent' because different communities of ideas *are not incommensurable*, as the relativist philosopher would like us to believe. As Searle once noted, the Inuits' historical origins as told by anthropologists (crossing the Bering Strait circa 13,000 years ago) is nearer to the truth than the Inuits' own creation myth (thrown out of a great crater that opened up in the earth…). And this is obvious to anyone who knows both belief-systems, just as it would be to an Inuit who had studied anthropology at Harvard.

[16] Popper treated absolute truth as a directive concept in Chapter 10 of his *Conjectures and Refutations.* Kant originated the view that there are directive concepts which lack a possible basis in our experience, but are still able to perform the pragmatic function of guiding our intellect in the direction of further syntheses. This was the case of his *ideas of reason*. According to the *Critique of Pure Reason*, they are concepts that reason uses in its striving to unify our knowledge, though unable to find satisfaction in sensory intuitions (1787, A 484, B 612).

conjunctions of propositions belonging to it do not generate a contradiction. Consistency may be a necessary condition for coherence, but it is surely not sufficient. For instance, consider the elements of the consistent set {Shakespeare was a playwright, lead is a heavy metal, 7 + 5 = 12}. They do not contradict one another. But since they do not have anything in common, taken together the elements of this set increase neither the coherence nor the truth of its elements; and we could create a set of this kind as large as we wish with 'zero' coherence. Consistency may be a necessary, but it is not a sufficient condition of truth. And worst than this is when we perceive that any definition of truth based on consistency alone would be circular, since consistency, being defined as the absence of contradictions generated by the elements of a set of propositions, assumes that their conjunctions *cannot be false*, in this way requiring the concept of truth-value in its own *definiens*.

More than just being consistent, coherence must be defined as *inferential*. The coherence of a belief system, of a system of propositions, is in fact determined by the dependence of this system on the *inductive and/or deductive relationships among its propositions.* This means that the degree of coherence of a proposition *p* should be determined *by its inductive and/or deductive relationships with the system to which it belongs* (*Cf.* Bonjour 1985: 98-100). Indeed, we know it is true that Mary was breathing the whole night long, because this is inductively supported by everything we practically and scientifically know about human metabolism and behavior, and this is a truth concerning our system of reality.

However, if we consider coherence as the only and proper mechanism able to generate truth, this last definition also leads to circularity, since the concepts of inductive and deductive inference used in the *definiens* of coherence *are also defined by means of truth!* A strong inductive inference is defined as an argument (or reasoning) that makes a conclusion probably true, given the truth of its premises, while a valid deductive inference is defined as an argument (or reasoning) that makes its conclusion necessarily true, given the truth of its premises. Consequently, the coherence account of truth can only generate the truth of any proposition of the system by assuming the independent truth of at least some of its other propositions, which makes the coherence view clearly circular. Any form of pure coherence theory is the victim of a *petitio principii*, as it simply assumes what it aims to explain.

18. Coherence as mediator

The view of coherence that I wish to propose here enables us to circumvent the difficulty. The reason is that in my understanding, coherence must be seen as *a complementary dimension of adequation theory*, namely, the condition that enables the *transmission* of truth in a network of thought-contents, usually beginning with those that are based on empirical (sensory-perceptual) experiences and/or some assumed formal evidence/assumptions (axioms or postulates).

Such view allows us to accept some factual content that should make some proposition true without the need for reducing this factual content either to some corresponding formal axiom or to an obvious perceptual or self-psychic thought-content. For instance, we know that the statement 'Mary was breathing when she was asleep last night' is true, and it is true because it corresponds to the factual content that Mary was breathing during her sleep. But usually we reach our belief that such a statement is true by adequation to a fact, not by means of direct observation, but by means of coherence, that is, by means of inferences derived from our system of beliefs. These inferences transmit what we may call *veritative force* – which we may define as any probability of truth higher than 0.5 – from one proposition to another. However, this veritative force cannot arise from propositions without truth-value, but instead is derived from propositions whose truth-value is ultimately based on (in Mary's case) a myriad of past judgments. These correspond to perceptual experiences that are the ultimate sources of our knowledge of biological laws, as well as our common awareness that Mary is a living human being like us and subject to the same natural constraints.

We begin to see that even if coherence cannot be regarded as defining truth, it plays an important role *as a mediating procedure whereby adequation is an indispensable ground*. For example: the modal proof of $P \rightarrow \Diamond P$ in our formal example does not come directly from AS1 and AS3 plus some rules of propositional logic. We first take a series of deductive inferential steps, and these steps are already constitutive of a linear coherential dimension of the verification procedure, which some coherence theorists erroneously saw as the proper criterion of truth for the formal sciences. In this modal proof coherence is constituted by implications transmitting veritative force – here understood as material implications from logical-conceptual, self-verifying truths postulated as axioms – but, as already noted, inevitably containing inductive inferences in the case of the verification of empirical thought-contents.

19. Roles of empirical coherence

The trouble with the coherence of empirical truth can be better illustrated by examples able to make clearer the relationship between coherence and correspondence or adequation.

First, suppose that someone anonymously sent me a package per post. I open it and see that it contains a book called *The Cloven Viscount* by Italo Calvino. I wonder if a friend named Sylvia sent it to me. I once knew Sylvia as a literature student in Rome, and at that time I gave her a copy of Calvino's book *The Invisible Cities*. However, the package was mailed from Rio de Janeiro. Thus, I realize that this book could have been sent to me by someone else. But then, I remember that Sylvia told me that she was born and lived most of her life in Rio de Janeiro. Hence, she could well be back at home in Brazil. An advocate of the coherence theory of truth would say that the thought-content of the statement 'p,' understood as abbreviating 'My friend Sylvia sent me a copy of *The Cloven Viscount*,' is made true by its coherence with other thought-contents, which can be ordered in the following way:

1. I received as a present the book *The Cloven Viscount* by Calvino. (r1)
2. Sylvia was a literature student when I knew her in Rome. (r2)
3. I gave Sylvia as a present a copy of *Invisible Cities* by Calvino. (r3)
4. (from 1, 2, 3) The book could have been sent by Sylvia. (s)
5. But the book was mailed from Rio de Janeiro. (t)
6. (from 4, 5) The book wasn't sent by Sylvia. (u)
7. Sylvia told me she had lived most of her previous life in Rio de Janeiro. (v)
8. (2, 7) Sylvia finished her studies in Rome and returned to Rio de Janeiro. (w)
9. (1, 2, 3, 5, 8) My friend Sylvia sent me a copy of *The Cloven Viscount*. (q)

What we really have here is an indirect procedure by means of which adequation is verified via coherence. To see this better, we need only revise the above reasoning, rejecting the partial conclusion u because of v. As a result, I can build the following coherent set of beliefs: {$\underline{r_1}$, $\underline{r_2}$, $\underline{r_3}$, $\underline{t}$, $\underline{v}$, $\underline{w}$}. Together, these belief-contents inductively make the conclusion $\underline{q}$ very probable. This anterograde set of reinforcing premises makes me – starting with the guess '?p' ('Was it Sylvia who sent me the book?') – see the identity of thought-contents p = $\underline{q}$ and conclude with practical certainty

$\vdash p$, affirming that it was Sylvia who sent me Calvino's book. However, each element of the coherent set of beliefs $\{\underline{r}_1, \underline{r}_2, \underline{r}_3, \underline{t}, \underline{v}, \underline{w}\}$ has its truth directly or indirectly based on correspondence.

To sum up, I agree with Stephen Walker's argument that a pure coherence theory is impossible (1989). Coherence could only exist independently of adequation if we were able to assume that e-thoughts could acquire probability or formal certainty independently of any anchorage in sensory-perceptual/self-sensory experience or in the axioms or postulates of a formal system. But, as our examination of the nature of coherence has shown, this would be circular. Moreover, consider again the example offered above. The thought-contents expressed by the statements that by means of coherence make the correspondence between p and $\underline{q}$ probable are all in some way observationally anchored. Either they describe a perceptual thought ('I knew her in Rome,' 'I gave her a book…,') or report testimonial information ('She told me she lived all her earlier life in Rio') or describe a personal experience ('I read the book…') or an inference ('She may be back home in Rio…') from testimony ('She told me…') based again on the sensory experience of others.

What was given to me as a *fact* in the above example was an *indirect* product of correspondences of other thought-contents with their own factual contents. And the increase of the veritative forces resulting from these experiences is what inductively warrants $\underline{q}$ to me as the derived proposition representing the fact that Sylvia sent me the book. The assumed warrant of $\underline{q}$, in turn, is what makes the e-thought-content of p true for me. In summarized form, introducing the symbol '~>' to represent strong inductive and/or deductive inference, the anterograde reasoning that leads to this attribution of truth can be symbolized as:

$$?p, \{\underline{r}_1, \underline{r}_2, \underline{r}_3, \underline{t}, \underline{v}, \underline{w}\} \sim> !\underline{q}, p = \underline{q}, / \vdash p$$

This helps us to understand better how coherence plays a role in the truth-discovery process. And it shows us why the coherence of our claims would have no force if it weren't anchored in perceptual experiences taken as evidence in the case of empirical truths, and in axioms or postulates in the case of formal truths. This is also why a fictional text can be perfectly coherent without in this way representing any factual truth concerning the real world: its anchors are only imaginary ones.

This kind of reasoning invites us to think that adequation comes first, since this kind of correspondence is what reveals truth. Moreover, in cases like, say, sensory-perceptual knowledge, we can in a sense have correspondence without coherence, while there is no coherence without

correspondence. However, this conclusion can be considered simplistic for the following reason. Correspondence without coherence must be impossible because of the fact emphasized by philosophers of science that *all observation is conceptually charged* or *theory impregnated* (Duhem, 1906, Ch. 6, sec. II; Popper, 1972, Ch. 2, sec. 18). In order to be conceptualized, experience already requires coherence with at least one sub-domain of our belief-system.

Nonetheless, I think that I can give a stronger justification for the indispensability of correspondence as the origin of veritative force by considering the real origins of the own input that a particular sensory-perceptual observation receives from our belief system. Suppose you go for a walk in a beautiful nearby field and you cannot believe what you see there: you think you are seeing a live unicorn! Soon you will begin to distrust your own senses, since you have learned that unicorns do not exist. Later the mystery is solved. You hear that it was actually a fake unicorn: a film production team had attached a horn to the forehead of a small white horse to create the illusion of a real live unicorn. Between scenes, the make-believe unicorn is allowed to graze in the field. The defender of coherence theory would say this proves that even sensory-perceptual observation can be falsified by our system of beliefs alone. But this argument is completely refuted when we consider that what was really responsible for your mistrust was not our system of beliefs alone, but the adequation of other perceptual experiences belonging to this same system or sub-system of beliefs. Indeed, we all know that unicorns are mythological creatures, and there have been no scientifically confirmed observations of unicorns or their physical remains, such as bones, fossils, tissue, etc. Nor have we found depictions of unicorns in cave paintings from prehistoric times, while we have found paintings of aurochs, for example. Moreover, we also know that evolutionary classifications of animals like horses and goats rule out the possible existence of unicorns. But these firm convictions against the existence of unicorns were all reached with the aid of induction by means of a multiplicity of other testimonial sensory-perceptual observations that were historically and scientifically made and passed on to us! This means that your sensory-perceptual observation of a unicorn was in the end discredited not by your system of beliefs independently of adequation, but by counter-evidence derived from the veritative force of other beliefs, all of them anchored in their proper adequation to perceptual observation.

Now, suppose we call '!u' the factual statement 'I am looking at a unicorn' and '~u,' its denial, based on the firm belief that there are no unicorns, which is grounded on the accepted zoological system of beliefs

that is in its essentials based on a multiplicity of observational experiences ‘e,’ questioning the possibility of !u, and we call ‘i’ the supplementary information given to you regarding the make-believe unicorn. We can symbolize the procedure that leads you to conclude the obvious falsity of u in two steps that jointly form a retroanterograde verification procedure:

(1) !u, (e ~> ~u), ~u, u ≠ ~u / ├ ?~u,
(2) ?~u, i ~> ~u , ~u = ~u / ├ ~u

Putting my argument in other terms: I certainly agree that sensory-perception is the immediate origin of the veritative force of a perceptual judgment, and this judgment can gain or lose veritative force due to greater or lesser coherence with our system of beliefs. However, this confirming or rejecting coherence acquires its *own* veritative force only by means of other sensory-perceptual observations whose truth is based on adequation. And reflection on this leads us to the inevitable conclusion that in one way or another the real ultimate origin of the veritative force of empirical judgments is always sense-perception, giving coherence the secondary, even if indispensable, role of transmitting the veritative force gained by means of sensory-perceptual experiences of adequation. My conclusion is that under closer scrutiny the supposed counter-example shows that correspondence comes first, simply because it is the only real source of truth. Thus, instead of defending an impure coherence theory, as Walker endeavored to do, I defend what he would probably classify as an ‘impure’ adequation theory – what I more accurately prefer to call an adequation theory that *incorporates* coherence.

20. Reverend David’s case

To reinforce my point, I now offer a second, more distinctive empirical example of the incorporation of coherence in correspondence/adequation. It concerns a judge’s verdict. It is well known that court rulings in criminal trials frequently cannot rely on direct perceptual evidence supplied by witnesses. Because of this, they are often heavily dependent on coherence, on proof by means of circumstantial evidence. This was the case with an American minister named Reverend David, who shortly after marrying a certain Mrs. Rose was admitted to a hospital suffering from severe abdominal pain. Since examination showed a high level of arsenic in Reverend David’s blood, a thought-content that we abbreviate as ‘!r,’ the following suspicion arose as the result of abductive reasoning: ‘Did Mrs.

Rose try to poison Reverend David?' in short, '?p.' The following additional factual evidence later confirmed this suspicion:

> s: Mrs. Rose had the habit of preparing bowls of soup for her husband, even bringing them to him in the hospital.
> t: Traces of arsenic were found in the pantry of Mrs. Rose's house.
> u: The bodies of Mrs. Rose's first three husbands, who all died of unknown causes, were exhumed, and it was not so surprising that high levels of arsenic were found in their hair.

We can now construct the following retroanterograde verification procedure:

$$!r \sim\!> ?p,\ \{!r \,\&\, !s \,\&\, !t \,\&\, !u\} \sim\!> !q,\ p = q,\ /\!\vdash p$$

Certainly, the conjunction of the statements r, s, t, and u gives us a strong inductive inference assuring us practical certainty that !q, which states an unobserved dynamic fact (namely, that Mrs. Rose did indeed try to poison her husband). This inferred factual content confirms our initial suspicion ?p derived from !r. However, a crucial point to be noticed is that factual statements r, s, t, and u are all considered true either by direct adequation with public factual observation or by derivation from publicly observable perceptual factual contents. Again, what is shown is that the element of coherence cannot stand alone. The plausibility of q is grounded on the conjunction of the observational statements r, s, t and u by means of coherence. But these statements are all true because of their direct or indirect adequation with perceptual contents, even if they may also rest on empirically grounded theoretical assumptions, the latter in some way also derived from other perceptual experiences. As we see, coherence alone cannot prove truth, because inductive and deductive coherence relations are ways of preserving and not of finding truth.

The conclusion is the same: coherence relations work like the high voltage power lines of an electrical power grid: though they are not able to generate electricity, they are able to transmit it over long distances. A plausible coherent system is not an independent mechanism, but only an inferential network over which the truth arrived at by means of originary adequation is transmitted. In other words: coherence only transfers the veritative force generated by the adequation of the contents of more basic beliefs concerning empirical or formal facts to derived beliefs or thought-contents. This transference of veritative force within a belief-system can act to produce an e-thought-rule that we believe corresponds to a non-

observed fact, which in my present example is q: the attempted murder using poison. The thought-content p is accepted by us as representing the factual content q, because both have the same content (structural isomorphism, etc.) which makes p true. Because in various ways q is reinforced in its application, we accept it as factual evidence of p's truth. And statement p is true because it *corresponds* to the fact that Mrs. Rose poisoned her husband, Reverend David, even if we know this fact not by observation, but only *indirectly*, from its coherence with other thought-contents that are observational and match their facts in a direct way. The thought-content q, the truthmaker of p, as I intend to explain, has a kind of Janus face: on the one hand, it expresses here a basal thought-content (an e-thought-rule or proposition), and on the other hand, it represents what we by indirect means are sure is an objective factual content, namely, the *fact* that Mrs. Rose tried to poison Reverend David. All this shows that coeherence is nothing but an interdoxal mechanism by means of which adequation can transfer its veritative force. It is by this means that coherence helps in confirming the truth of statements.

Now, concerning the truth of the observational statements r, s, t, u, we return to the point already made when we analyzed our first example. Each of these observations is embedded in at least some subsystem of beliefs. Although a given observation r makes its own contribution to truth by means of direct adequation with a fact (the high level of arsenic in the blood), it can be reinforced by its coherence with the accepted subsystem of beliefs in which it is embedded (like s, t, u together with the hypothesis p), or even be refuted by other beliefs of this same system. But here again, the consideration of this network of giving and taking among sensory-perceptual and derivative beliefs leaves no room for a veritative force arising from coherence.

The important question that remains open is about the precise status of the statements of factual evidence (like of q) in our examples. It is the expression of an e-thought-content-rule, but it must also be seen as able to represent the actual factual content, namely, a cognitively independent external criterial tropical arrangement. Are these two possibilities reconcilable?[17] This crucial question will be tackled in the following sections.

[17] If *q* were only the direct expression of a factual content, we would fall into a kind of *strong externalism* that admits that part of our content-thought-meaning is a directly given fact in the world (a 'structured proposition' or something of the kind). However, without further qualification this view would demand too much from our epistemic powers, leaving unexplained not only the possibility of falsity, but also the inevitable fallibility of our supposed knowledge of truth.

21. What about the truth of the truthmaker?

One of the most serious problems for the adequation theory of truth concerns the infinite regress that arises from factual evidence that verifies suppositions, that is, verifiers or truthmakers. We can pose the problem in the form of a dilemma: Either the truthmaker – the evidential fact, the real or actual factual content – is unquestionable, or it can be doubted. Suppose (a) that the evidential fact is unquestionably true. In this case, we seem to be guilty of dogmatism, because we treat our normal perceptual truths and even purely self-sensory truths[18] as if they were beyond any possibility of being false. But this would be *to deny the fallibility* of sensory-perceptual knowledge. We cannot be absolutely certain about the evidence for any (or maybe almost any) empirically given factual content. Even formal axioms always have a degree of arbitrariness in their choice and can lose their applicability after changes in our broader system of reality. Now, suppose (b) that we consider the evidential content believed to be a fact (which shows itself as a thought-content) as open to doubt. In this case, it seems that we need to search for new evidential content (another thought-content) that would warrant its truth. Since this new factual content will likewise not be beyond doubt, we would have to look for further evidential content and so on endlessly. Since we cannot stop this regress, we have no way to ground our suppositions, because any ground we find will lack the necessary solidity. The upshot is that neither alternative (a) nor alternative (b) is satisfactory.

Restricting myself here to the cases of external empirical truths, I think we can solve the dilemma if we consider examples in sufficient detail.[19] Consider the following example of an observational statement !o: 'There's a dolphin swimming in the sea.' Imagine that the truth of this sentence depends on the observation of a dolphin surfacing from time to time – an observation that can be interpersonally shared. For the first person who sees the dolphin, the procedure has a retrograde form:

$$!o,\ ?p,\ o = p \,/\, \vdash p$$

For a second person, already informed by the first and trying to locate the dolphin in the sea, it will have a retroanterograde form:

---

[18] When I write of purely sensory truths, I am thinking of cases covering false sensations and feelings, such as imaginary pain induced by hypnosis or an emotion that someone defensively substitutes for the true one.

[19] A deeper understanding will demand a response to the problem of perception that will be attempted later in this chapter.

$p \sim\!> ?p, !\underline{o}, p = \underline{o} / \vdash p$

But this does not mean that !o, the given evidence, is absolutely warranted! It can be *defeated*. Suppose that due to a scarcity of real dolphins and in order to entertain tourists, a diver is hired who swims just below the surface with a rubber dolphin mounted on his back, surfacing from time to time in a way that gives dolphin watchers the illusion that they are seeing a real dolphin.[20] In face of this, the factual content !o that should ground the verification of ?p is defeated. Those aware of the deception could correctly point out: 'It is false that there is a dolphin swimming in the sea.'

However, it should not be hard to find a solution to the problem. What we believe to be factual content need not be regarded as absolute. It can be seen as a thought-content *assumed* to unquestionably represent an actual factual content (the ultimate truthmaker) *within the context of a practice that typically assumes that we do not have atypical circumstances that if present would defeat the assumption.* Thus, consider the linguistic practice (A), in which we recognize things in normal daylight that are large enough and near enough to be identified as dolphins, and they are employed in the context of a tourist beach where people expect to see dolphins swimming in the water offshore… In this practice we are allowed to *assume* that the observational content 'I am watching a dolphin that has just emerged from the sea' can be taken as *unquestionable evidence* expressible by *!o.* It is thereby a real-actual fact, a truthmaker or verifier that we accept as giving practical certainty to the thought-content that there is a dolphin in the sea near where the observer is standing. Assuming the information content and the context at our disposal in this practice, and assuming that all other things remain the same, seeing a dolphin must undoubtedly be accepted as the truthmaker of the hypothesis *?p*. Assuming that *o* also has internal phenomenal content (with psychologically given sensory impressions), we could say that in this case we are allowed to assume that the e-thought-content-rule of *o*, that is, *o* without the underline (expressible as: 'I am having visual impressions of a dolphin emerging from the sea') can be considered the vehicle of the experience of the real-actual fact *o* given in the world (representable as: 'Being a real dolphin that has just emerged from the sea'). Summarizing: in practice, our willingness to accept evidence is dependent on a *ceteris paribus*, namely, *on the assumption that the observation isn't being defeated by some condition extraneous to all that is expected for the working of the given practice*.

[20] I read this story many years ago, although I am unable to find the source.

Now, in the given case there is a defeating extraneous condition, which begins with the scarcity of real dolphins in the vicinity and ends in the training of a diver to swim just below the surface with a rubber dolphin mounted on his back, sometimes rising to the surface in a way that gives people on the shore impressions of seeing a real dolphin… Assuming that some observer S is aware of this information, what is given to him isn't the practice (A) but a different observational practice that we can call (B), which includes information about the very unusual background circumstances. In this (B) practice, we cannot postulate the observation of a real dolphin merely because we see what appears to be a dolphin emerging from the sea. Under the circumstances presented by (B), in which a rubber dolphin is often carried on the back of a diver swimming just below the surface, to know with certainty that one is observing a real dolphin would require closer and far more careful examination. Closer underwater inspection, for instance, might reveal factual evidence of a fake rubber dolphin, which can be symbolized by o'. In this new practice, the thought-content expressed by p could not be verified by the fact able to be represented by !o, because !o isn't really given to S, since we already know that in its context !o cannot be trusted to be a real dolphin. However, ?p could be falsified by the more careful observation provided by o', as the following retroanterograde schema shows:

$$p \sim\!> ?p, !o', p \neq o' / \vdash \sim p$$

What this example shows is that our usual certainty regarding experienced factual content, despite not being absolute, must be *postulated as certain or irrefutable!* This is *assumed* as a practical certainty and must be treated as beyond the level of a merely probable truth, under the assumption that the factual context does not involve unknown evidence able to *defeat* the linguistic practice in the context of which the perceptual judgment is made. If we obtain information indicating different background circumstances able to discredit the practice sustaining the perceptual judgment, as in the case above, the assumed evidence vanishes.

I can offer a second, similar example, only to reinforce the point. Yvonne is driving a car through a desert, and she thinks she sees a lake, but it is really only a mirage. At first, she believes the lake she sees on the horizon is real. We can symbolize this through the following retrograde verification procedure:

$$!o, ?p, o = p / \vdash p$$

However, it soon becomes clear to her that she has made a naïve mistake; what she really sees is nothing but a so-called inferior mirage. This is caused by the refraction of sunlight passing through a layer of hot air near the ground. In this way, she adds to the background conditions the easy graspable unusual circumstances able to invalidate normal perceptual evidence. As she has learned that these unusual circumstances defeat the rules of normal observational practice (A). Instead of thinking !p, 'I see a lake', she thinks ├ ~p 'I do not see a lake,' eventually concluding: ├ q, which asserts the sentence 'I see an inferior mirage' (or 'I see the refracted blue of the sky'), which represent a different factual content that can be represented as o.'[21] Consequently, what was at first accepted as external evidence is now viewed as an erroneous interpretation of phenomenally given data, since practice (A) was replaced by the new practice (B). The gained awareness of the context allows the invalidation and replacement of what was at first assumed as an unassailable truthmaker. We can symbolize this change through a sequence of the two following anterograde verification procedures belonging to practice (B):

?p, !o', p ≠ !o' ├ ~p,
?q, !o', q = !o' ├ q

It is worth noting that in both interpretations the phenomenal content of perception remains the same: an impression of seeing the color blue near the horizon. But the interpretation of this content is very different, once o' is read as a new factual content: a mirage existing in the world. And Yvonne understands what she sees differently because a more complete awareness of the background information given by the surrounding circumstances (including the fact that the blue band always keeps the same distance to the car) is able to defeat the seemingly reasonable initial interpretation of the visually-given content as o.

22. Objection of the linguistic-cognitive circle

Probably the most influential epistemic objection to the correspondence theory of truth is the so-called problem of the *linguistic-cognitive circle*: Propositions can only be compared with propositions. If we compare hypothetical propositions with propositions representing evidential

[21] Even though the phenomenal contents of o and o' are similar, the whole factual context must be very different, since at least the dispositional properties of 'the blue there' must be completely different.

contents, even if these are taken as irrefutable, we remain trapped in our language and thought. Even if we find the strongest factual evidence, this evidence could only be considered in the form of linguistic expressions of propositions, but in no way do we find evidence by direct comparison of propositions (even if understood, as we do, as e-thought-rules) with real facts, states of affairs or events in the world (Neurath 1931: 541; Hempel 1935: 50-51). Here again, we would be in danger of ending up in an infinite regress with epistemic skepticism as a corollary.

A *prima facie* general reply to this objection is that saying we are trapped in an intra-linguistic or intra-cognitive world already assumes we know there exists an extra-linguistic and extra-cognitive external world – a knowledge that remains unexplained.

Philosophers like Moritz Schlick (1936) and A. J. Ayer presented a more focused reply. Here is A. J. Ayer's well-known reply:

> We break the circle by using our senses, by actually making the observations as a result of which we accept one statement and reject another. Of course, we use language to describe these observations. Facts do not figure in discourse except as true statements. But how could it be expected that they should? (1963: 186)

Ayer's argument contains a strong appeal to common sense. Nevertheless, this appeal seems to contradict another enduring idea, which is also not alien to common sense. It is the idea that the whole content of our usual perceptual experience should be some kind of conceptually articulated belief-content and therefore should be *mental in nature*. Consequently, it remains not entirely unreasonable to think that we could never have direct and unquestionable access to anything referred to by a perceptual thought, even if considered as e-thought-rules, namely, external facts as they are in themselves (*Cf.* Blanchard 1939, vol. 2: 228).

One reaction to this dilemma would be to accept the kind of last resort solution called *idealism* (e.g., Foster 2000). But today idealism seems to be an almost forbidden solution. According to idealism, all reality is in some sense mental. This view conflicts with one of our chief modest commonsense principles, namely, that we are surrounded by a cognitively independent external material world. In fact, our empirical knowledge (particularly our scientific knowledge) has told us that the mental is in some sense a minuscule emergent portion of the physical world, dependent on it to exist, just as the phenotype is dependent on the genotype. In other words, the mental appears to supervene the physical insofar as experience – scientific or otherwise – has shown. Moreover, if we stay on the side of our principle of established knowledge (Ch. II, sec. 5), idealism will

remain anathema, since it denies not only the modest commonsense truth that the external world is non-mental, but also the scientific truth that the external world as a whole is overwhelmingly non-mental. In some non-mystical sense of the word 'emergent,' science has shown that mind is an emergent property of life, which is an emergent property of organic chemistry, a rare carbon-based chemistry emergent from our atomic and sub-atomic physical world. And all our astronomical knowledge conspires to show that this minuscule accidental phenomenon of the emergence of the mental is destined to disappear with the unavoidable process of death of the universe, which is foreseen by the laws of thermodynamics. Finally, from an anthropological perspective, idealism is very often motivated by *wishful thinking*, as is argued in the philosophy of culture and the humanities by authors ranging from Nietzsche to Freud and from Hume to Marx and Durkheim. It seems that human beings pay a high price for having acquired consciousness. In some way, it recalls the price paid by Prometheus for his theft of fire to benefit Mankind. Even if consciousness makes us better able to survive, it also gives us an increasing awareness that we live in an unpredictable and dangerous world, along with a clear sense of our own physical vulnerability and finitude. Idealism, by making the external world in some way mind-dependent, can be helpful in supporting those illusions of control over the external world that could give us some hope of beating the odds, a thought that is made explicit in Berkeley's writings. Summing up, due to all the knowledge we have at our disposal today about the physical world and ourselves, more than ever before we have strong *external* reasons to reject idealism in favor of epistemic realism. (The internal reason is what I intend to expose later.)

23. Answering the objection of the linguistic-cognitive circle

Epistemic realism concerning the external world can be understood as the view that preserves the natural opposition between the mental and the material worlds in the sense that we can roughly characterize the internal mental world as only experienceable in the first-person, while the external physico-material world can be mainly characterized as able to be experienced in the third-person.

Assuming epistemic realism, in what follows I will defend *direct realism* as able to give us the kind of epistemological framework that will make it possible to break the linguistic-cognitive circle. Direct realism is the view that our senses provide direct awareness of the external world, showing it pretty much as it is. Direct realism differs from *indirect* or

representational realism, which is the view that we have direct experience only of our own sensations, which inform us about the external world, so that the latter is never directly experienced. Both, direct and indirect realisms, differ from a third traditional epistemological position, called *phenomenalism*. According to this last view, we can have experiential access only to our sensations or sense-data, since there is no sufficient reason to postulate an external world independent of actual or possible sensations. This view leads almost inevitably to idealism and to rejection of a really existing non-mental external world (*Cf.* Ch. IV, sec. 20).

My defense of direct realism begins with the suggestion that everything experienced in real perception has a kind of Janus face, able to explain the double nature of *!o*, as the thought *o* and as a fact in the world underlining *o*. What I mean is that what is given to us in proper sensory-perceptual experience of the external world can always be understood as two different types of interrelated entities: one psychological and the other physical, as follows:

> (A) The *merely psychological experience* of cognitively-dependent internally given *sensory content*, the so-called *sense-data*.
>
> (B) The *proper, physically understood* cognitively-independent, externally given *perceived content* (that is, the external real entities understood as physically particularized property-tropes, material objects as clusters of tropes, simple or complex facts as tropical arrangements…).

Psychological experience (A) gives us what we may call *sensory impressions or sensory contents* (also called sensations, sensa, sense-data, percepts, phenomena, representations, ideas…). It seems commonsensical that sensory contents are always present in perceptual internal tropical experience (even if we are usually unaware of them) as I intend to show later. But experience (B) also seems beyond doubt: it is the view that in addition to sensory experience, when we really perceive something, this something is given to us as an external, physico-material kind of entity. Indeed, it is also commonsense knowledge to say that we usually perceive the external world *directly and as it really is*. And this external world, as we have shown, is originarily accessible as constituted by physical, external tropes (properties) relatively dependent of clusters of relatively independent compresent external tropes with some form and mass, most of them called material objects, and by arrangements of both, also called facts.

The clearest evidence favoring this double view is given by tactile experience. Suppose I touch a hot stove with my hand. I can say *I have a sensation of heat*: this sensory-impression is the psychological (criterial) sensory-content of experience (A). Alternatively, I can also say that *I have perceived that the stove is hot;* this is the correct perceptual experience of the (criterial) perceptual content, that is, an externally given physical tropical state of a material object (B). The most important point is that in the normal case we cannot *phenomenally* and *descriptively* distinguish experience (A) from experience (B) (*Cf.* Searle 2015: 24). In spite of this, we can always conceptually distinguish the two cases, as the following examples of tactile experience show:

(A) [I have the feeling that] the stove is hot.
(B) The stove [I am touching] is hot.

In a similar way, I can say:

(A) [I have the feeling that] I am holding a tennis ball in my hand.
(B) I [am aware that] I am holding a tennis ball in my hand.

Now, from auditory experience, I can say:

(A) I [have the auditory impression that] I hear thunder.
(B) I hear thunder [outside and over there].

And of the most common visual experience, I can also say:

(A)[I have the visual impression that] I am watching a fishing boat entering the mouth of Pirangi River.
(B)[I am aware that] I am watching a fishing boat entering the mouth of Pirangi River.

As you can see, although what we could call *linguistic descriptions of contents* outside the brackets *are the same* in cases (A) and (B),[22] in (A) cases I speak of *merely sensory (criterial) contents occurring in my head*, while in (B) cases I speak of *objectively real physico-material external contents – perceived factual* (*independent criterial*) *contents pre-existing*

[22] Searle uses the expression 'phenomenal appearance,' but then we should distinguish the *psychological* phenomenal appearance from its correlative *physical* phenomenal appearance.

in the external world. Note that in cases of perceptual contents, I speak of contents such as the distinguishable objects found in a drawer, that is, of objectively real tropical entities given to experience, which should not be confused with *semantic* contents understood as rules whose dependent criteria should be satisfied by the first ones). Furthermore, on the one hand, the real perceptual content (B) is *epistemically dependent* on mere sensory content (A), because without sense impressions (A), one couldn't know (B); on the other hand, sensory content (A) is *ontologically dependent* on the real external things constituting perceptual content (B), since (B) causes (A).

Accepting the above dual understanding of perceptual experience is not hard and does not compromise direct realism. I can illustrate how harmless the duplicity is by comparing it with our interpretation of objects that I see in a mirror. What I see in a mirror can be understood as: (A') a simple image of things, for instance, the image of a vase of flowers on a table. But it can also be understood as (B') the vase in itself that I am looking at in a mirror. For instance, I can point to the object I see in a mirror, and you can ask me if I am pointing to the reflected image of the vase of flowers or to the real vase of flowers. That they belong to different domains of experience is made clear by contextual differences: the image isn't considered real, because I cannot touch or smell it. The real vase of flowers, on the other hand, can be touched, smelled, directly seen from all sides, manipulated, broken; its weight and its size can be accurately measured and shown to remain constant, independently of the changeable apparent size of its image… Alternatively, I can change the apparent size of the image by bringing the vase closer to the mirror. And this apparent size always doubles the real distance of the vase from the mirror… Nevertheless, to a reasonable extent, qualitative properties and relations of both image and reality will be alike or correlated. Moreover and unavoidably, looking in the mirror I would not be able to see and locate the vase on the table *without the help of its image.*

In fact, access to the real vase is dependent on access to its image. As in cases like (B) above, (B') is *epistemically* dependent on (A'), because without the image (A') I could not see (B'). Alternatively, (A') is *ontologically* (causally) dependent on (B'). This is why when I pay attention to an object in a mirror I interpret it as perceptually dependent on its image, but when I pay attention to the image I see it as causally dependent on the real object. I can easily say I see the reality *by means* of the image. But I will never say that I cannot see the actual object only because what I really see is just its image.

Like all analogies, the mirror-image analogy has its limits. For instance, I can always be aware of the image in the mirror as an image, but I am normally unaware of my own sense-data (except, for instance, in cases like those of lucid dreams). However, even here we find something similar: I am aware of the image *qua* image externally, mainly through conditions like the restriction to visual access and the relations to other things, not due to the image itself. Anyway, the mirror-analogy reinforces the idea that we can answer the objection of the linguistic-cognitive circle by saying that the content of any real experience can be understood in two ways:

(A) Internally and psychologically, as a first-person sensory-based e-thought-content-rule (a sensory-perceptual e-thought-content-rule with its internally fulfilled criteria).

(B) Externally as a third-person physico-material fact (the referred non-semantic factual content constituted by arrangements of external tropical criteria).

Now, *insofar as we are also able to read in the given phenomenal content an external factual content, we should be able to escape the linguistic-cognitive circle.*

A complementary but also indispensable point that I have dealt with many times already is that we almost never have a complete sensory-perceptual experience of external factual content. Our perceptual experience is typically perspectival. We experience only facets, aspects, sub-facts. If from a position on shore I see a fishing boat entering the mouth of Pirangi River, I may experience (see) only one side of the fishing boat. However, based on this dynamic tropical sub-fact (an aspect of a process), I am able to say not only that I see one *side* of the boat – the *sub-fact* – but also that I *see the whole boat* and that I am following *the whole process* of the real fishing boat entering the mouth of Pirangi River – a *dynamic grounding fact* (See Ch. IV, sec. 25-27; Ch. VI, sec. 6). All these descriptions might be true and their truth derives equally from adequation.

Another complementary point is the unavoidable admission that sensory content (sense-data) really accompany all our perceptions. That this purely sensory content exists can be illustrated by a phantom pain from a missing limb, after-images, and lucid dreams. A person can feel pain in an amputated limb as if the limb were still there. An after-image appears when someone closes his eyes after looking briefly at the sun. A

lucid dream is a dream controlled by a person who is aware that she is dreaming. Furthermore, for those still skeptical of the existence of internal sense-data in normal perception, experiments with vision reconstruction, which involve computationally reconstructed brain experiences of scanned moving images by means of fMRI (e.g., Nishimoto et al. 2011), are more than proof that these sensory contents in the brain really exist, as in these experiments subjects experience their own sensory images and interpersonally compare them with what they see in the external world![23] The dichotomy considered above is also important because it is a necessary condition for the already noted defeasibility of observational evidence: under perceived anomalous conditions we can reinterpret experience by withdrawing from what we believed to be real perceptual content to mere sensory content reinterpreting the lost one.

24. The argument of illusion

Against the kind of direct realism explicated above and favoring indirect realism or even idealism, there are two well-known traditional arguments: the *argument of illusion* and the *argument of science*. As almost too much has been written against these arguments,[24] I will emphasize only the essentials. I think that answering these arguments strengthens my own moderate direct realist view.

I begin with the argument of illusion. It usually concerns cases of perceptual illusions in which we seem to perceive something that should not be perceived, particularly in the extreme case of hallucinations. There are many examples that support this argument. They all aim to prove that in the best case perception is *indirect*, since it always occurs through the 'veil of sensations.' In what follows, I summarily present several examples, some of which were already known in antiquity:

1. I go outside in mid-winter without wearing gloves, although the temperature is minus 26 degrees. When I come back inside, my hands are stiff from the cold and I cannot feel them. I soak my hands in water that is only at room temperature and yet feels warm! Generalizing, *what I directly feel are my sensations*, and only through them can I gain information about external temperatures…

[23] It is true that fMRI measures brain activity by detecting changes in blood flow, but blood flow and neuronal activation are coupled.

[24] For an admirably intelligent and vivid defense of direct realism, rejecting the argument of illusion, see John Searle 2015.

2. I am near a speedway. A car passes me driving at a very high speed. Because of the Doppler Effect, its sound changes pitch from high to low. Hence, I do not hear the true sound, but *only experience my own auditory perception*, which gives me information about external sounds.
3. A person with jaundice may in some rare cases see the world as yellow due to an accumulation of bilirubin in his eyes. Now, what allows us to claim that people who do not have jaundice see the world as it really is, in its *true colors*?
4. If I press the side of my right eye with my right finger, I have the impression that things in front of me are moving in the opposite direction. Since these things are not moving, I conclude that *I can directly see only my images of things*, that is, my sensory impressions, my sense-data, and not things as they are in themselves.
5. If I hold my index finger fifty centimeters from my face and focus on the far end of the room, I see *two* images of index fingers. If I then focus my eyes on the finger, the two images merge into a single one. Since they are not phenomenally different in the two cases, I conclude that *what I really see are only sensory impressions of my index finger,* even if I can secondarily locate my finger through these sensory impressions.
6. I look at a coin I am holding at an angle to my line of vision. I am convinced it is round, even though it appears elliptical. Indeed, only occasionally do I see a coin in what I consider to be its true round shape. Hence, what I primarily experience are *my own sensory impressions of elliptical forms that I think of as different views of its true round form.*
7. I walk around a table looking at it from different perspectives. Then I look at the same table from different distances. The visual impressions are always different. Consequently, what I see is not the table, *but only my own changing visual impressions.*
8. I see a lake in the desert, but soon I perceive that it is an inferior mirage caused by layers of hot air above the sand which refract the blue light from the sky. My visual impressions of a lake and a mirage are phenomenally the same, hence *what I primarily see are my visual sensory impressions* of a blue lake that is not really there.
9. Suppose I have a perfect hallucination of a white horse. What I see is not a real white horse, but only a hallucinatory image. Since this image made up of sense-data isn't different from what I see when I

> see a real white horse*, the primary object of perception must be my sensory impressions or sense-data.*

If the argument of illusion applies to cases (1) to (9), why not to all cases? Why not, as Bertrand Russell once suggested, be democratic and admit that in all cases we first need to perceive our sensory contents – the sense-data – in order to get information about the external world?

The conclusion suggested by the argument of illusion seems to refute direct realism, which should then be replaced by indirect realism – a view already accepted by Descartes and mainly attributed to Locke. The suggestion is that the objectively real world is always perceived *indirectly* through the *veil of sensations*, which is formed by sensory impressions or sense-data. To this one could add, using a Kantian argument, that we experience how external things are *for us* and never how they really are in themselves.[25] Nonetheless, against this one could also sustain that, since what external things are for us is the only way to tell meaningfully what they can be in themselves, what they are for us must also in some way be what they really are in themselves.[26]

25 Answering the argument of illusion

In my understanding of direct realism, I do not wish to deny that there are sensory impressions or sense-data; I do not even wish to deny that we perceive the world *by means* of a veil of sensations formed by sensory impressions, since by accepting (A) I accept these assumptions. What I reject is the claim that these things make our perception *indirect.* For as is well known, we never say we *perceive* our sensations; what we might say is only that we normally perceive the external world directly through our sensations or sensory impressions. This means that just because we can show that we perceive the external world by means of one or even several veils of sensations doesn't make our perception of the external world indirect, since it is a category mistake to defend this view. Put simply: the main problem with the argument of illusion can be seen as resulting from a

[25] Kant defended a minimalist form of direct realism: all that we are able to know is a multiplicity of in some way cognitively dependent phenomena resting on an unknowable thing in itself (*Ding an sich*).

[26] There is no view from outside. One can use different conceptual schemas (or languages) to say what something is for us. For example, I can say that this computer is made of metallic and plastic pieces, and I can also say that it is made of atoms. The macrophysical and microphysical schemas are complementary ones, and they both inevitably explain how something is *for us.*

misunderstanding of the semantics of our concept of *directness*.[27] Consider the following four sentence pairs:

1. I saw the Sun directly, through my green glasses.
2. I saw the Sun indirectly, in a dark room projected on a screen by means of a telescope.

1. The trip was *direct* (the bus traveled directly through Germany from Constance to Munich, with a lunch stop of thirty minutes).
2. The trip was *indirect* (it started with a bus trip from Constance to Lindau, where passengers completed their journey on a direct train to Munich).

1. The bullet struck the victim *directly* (after piercing a windowpane).
2. The bullet struck the victim *indirectly* (after ricocheting off a wall).

These examples show that what makes some relations direct is not necessarily the fact that we cannot find *intermediaries* between the *relata* – they very often exist and are more than just one. Directness/indirectness is to a great extent a conventional distinction that depends on the relevance of the intermediaries for what we aim to consider.

In the case of perception, language conventions allow us to say that we perceive things around us *directly*, even if by means of a causal process involving a number of intermediaries. Because of this, there is nothing wrong in accepting the view that we perceive things directly by means of our percepts or sensory data or through a veil of sensations, just as much as there is nothing wrong in saying that a victim was struck *directly* by a bullet, even though it first had to pass through a windowpane.

Having this in mind, if we again consider the examples of the argument of illusion one by one, it becomes clear that perceiving things through sensory impressions does not mean that we must perceive them indirectly:

1. I soak my cold, stiff hands in water that feels *as if* is warm. I am fully aware, however, that the water is actually at room temperature, and although I perceive the temperature *directly*, I know my perception is deceptive. I know very well that if my

[27] For similar lines of defense, see Austin 1962, Ch. 2; Cornman 1975, Ch. 2 and 6; Dancy 1985, Ch. 10; Lowe 1992; Huemer 2001, Ch. VII. Huemer proposes that we should sharply distinguish the *object* of perception from its *vehicle*, and Lowe points out that the veil of sensations must be seen as a *bridge* or a *window* to the real world.

hands were not cold, I would feel the water in its room temperature. Then I would feel it in a non-deceptive way because the normal functioning of our perceptual organs is an expected condition for adequate perception.

2. I hear the car's motor *directly*, though in distorted ways. If I could drive alongside the car at the same speed, I would hear it in an undistorted way; I would hear it *directly* as it is assumed to really sound, that is, free from the distorting Doppler Effect.
3. A person can say, 'I see things *directly as if* they were yellow, though I know that isn't their true color,' because he knows he has jaundice. – What we call the true colors of things are by convention the colors I see under what are considered to be normal conditions. This presupposes the right distance from them, having normal vision, seeing things in adequate illumination with a neutral white balance, etc.
4. Even if I show by pressing my eye that I see things as if they were moving through my visual field, this does not mean that I am not seeing them directly. In fact, I can even say, 'I see external things *directly* and precisely as they are, though having a false impression that they are moving.'
5. In this example, as Searle has noted, one can instead say, 'I do not see two fingers… In fact, I am *directly* seeing my own index finger *as if* it were doubled.'
6. Concerning the form of the coin, it appears elliptical, but I can say that I *directly* see a round coin that only 'looks elliptical' because it is being held at an angle. – As A. J. Ayer pointed out, what we consider to be the true form or the real color is partially a matter of conventions (*Cf.* 1973, Ch. 4). Here we have the convention that the real form of a coin or a table is the form we see when we look down on them from above. In the same way, we have a convention that the real form of a mountain is the form we see when looking at it on the level of the base at a certain distance, but not an aerial view from above (e.g., the Matterhorn, the Sugarloaf). Based on conventions defining the perception of things as they are (our normal perception), we say that the real color of a tropical mountain is green, even if it may seem blue when viewed from a great distance...
7. In the case of the different sensory images of the table, you always assume that you are seeing a table that is always one and the same. This shows that the different perspectives and distances are only variations in the way *the same* table 'looks like.' And these

different perspectives and distances are said to be different ways in which you *directly* see the same table.[28]

8. In the case of mirages, I see what looks like a lake, but usually I can say that I am aware that what I really see is the image of the sky refracted by layers of hot air on and above the desert sand, and I say that I see this mirage *directly*.
9. Finally, in the case of a hallucination, it is simply incorrect to say that I *see* the content of my hallucination. As Searle emphasized, I only *believe* I see it, when in fact *there is nothing there to be seen*! Verbs like 'seeing,' 'perceiving,' 'being aware of' are here primarily related to factual, external content, and not to merely internal sensory content. Even if we agree that it is *by means of* sensory content that we have perceptions of things, this does not make our realism indirect. In a similar way, when we say that a bullet struck a victim after piercing a windowpane, we do not mean that the bullet struck the victim indirectly.

This kind of answer is not as new as it might seem. It was already present in the following comment by the direct realist philosopher Thomas Reid targeting his contemporary David Hume, almost three centuries ago:

> …visible appearances of objects are intended by the nature only as signs or indications and the mind passes constantly to the things signified without making the least reflection upon the signs or even perceiving that there is such a thing. It is in a way something similar that the sounds of a language after becoming familiar are overlooked and we attend only to the things signified by them. (Reid 1967: 135)

To most present-day philosophers, including myself, direct realism is the most proper answer, and we can see the persistence of competing doctrines as a testimony to how slow and uneven progress can be in philosophy.

Summarizing: we perceive things directly, even under misleading conditions like those of delusions. This justifies the direct realist view of whatever is given in perception. And this does not mean that we do not perceive the world inevitably *by means* of a veil of sensory impressions or sense-data, just as seeing an object in a mirror does not mean that we do not perceive the object as it is *by means* of its mirror-image. This justifies

[28] Searle's conclusion goes to the point: 'The whole discussion presupposes that I am actually seeing the table throughout, for there is no way that the table could continue to present to me different appearances from different points of view, if I were not actually seeing the table.' (2004: 273) See also Huemer 2001: 119-124.

our psychological interpretation (A) of a given content as based on sensory data, without forcing us to reject interpretation (B). This reinforces the idea that a phenomenal content can be interpreted as psychological (constituted by sense-data), but also as physical (constituted by material things, their tropical properties, etc.)

26. The argument of science and its answer

Finally, a word about the argument of science. According to this argument, perceptual experience depends on causal physical stimulation of distal neuronal cells that through synaptic activation ultimately lead to the stimulation of cortical regions in the brain, which produces in us an awareness of the objects of experience. Thus, our experience is in fact the experience of something occurring in our brain, which is nothing but the experience of sensory impressions. Consequently, our direct experience can only be of sensory impressions occurring in our brain. From this it follows that we cannot have direct experience of the world around us and that we cannot be sure that our contents of experience reflect how the external world really is or even provide a warrant of its existence. Worse yet, we may be led to the incredible conclusion that since our brains also belong to the external world, we cannot even be sure that our brains exist... All we can be sure of is that there are sensory impressions!

The answer to the argument of science is that there is nothing semantically wrong in saying that we *directly* experience things given in the external world, even if this experience demands the underlying work of complex neuronal structures as intermediary means. In the case of visual perception, we have simulacra of things seen, first in the projected image of the object causing the activation of photoreceptor cells in the retina and then in a corresponding activation of the striate cortex in the occipital region, which is analyzed by the visual-association cortex...

In my view, the decisive point is that the sentence ‘we directly see the objects’ belongs to the conceptual schema expressed by our natural language, while expressions like ‘by means of…’ or ‘through…’ used in the argument of science belong to a physical-neurobiological language concerning the underlying intermediating physical processes responsible for neuronal activation-patterns that in our natural psychological language we use to refer to direct experience of the world around us. Each language – each conceptual system – works well in its proper field, and each language has its own way of segmenting or not the process of perception. *Mixing* both languages is what leads to fallacious reasoning. In the present case, the fallacy arises when we use the semantics of the physical-

neurobiological language – which has discovered complex causal processes at the physical and neuronal levels – to deny the semantics of our natural language – which establishes a direct relation of seeing or being aware of things in the outside world. This confusion is again a clear case of equivocity (Ch. III, sec. 11).

Finally, as far as I know, what we call sense-data in the visual case has much to do with the activation of the striate cortex, since the stimulation of this region without the activation of photoreceptors in the retina is apt to produce hallucinatory phenomena (Teeple, Caplan, Stern 2009: 26-32). However, this fact alone does not make visual perception indirect, since it isn't captured by the semantic conventions governing what we are used to call the directly perceived objects around us; and it is this psychological-natural language that is responsible for what we understand with the ordinary word 'perceiving.'

27. Question: How do we warrant the perception of external content?

Even agreeing with all these commonsense arguments made to show that we are able to have *direct* access to entities belonging to the external world by means of the veil of perception, that is, by means of sensory contents or sense-data, the phenomenalist can still pose the question: why are you so sure that the externally given tropical entities that you say your semantic-cognitive rules apply to really belong to a non-mental physico-material world? After all, as we learned from our discussion of Berkeley's and Mill's phenomenalism (Ch. IV, sec. 20), it does not seem inconceivable that the objects that satisfy these rules are only actual or dispositional configurations of mental or psychological sensations… which seems to lead us back to idealism.

However, for the already given external reasons (related to what science and culture presently have to say or suggest about the world and ourselves), idealism seems to be far from a plausible option. In what follows, I expect to give *internal* epistemological reasons to think that idealism is a philosophically equivocal solution. This means I need to give reasons for our commonsense assumption that the external fact that we believe to satisfy the dependent criterial configurations demanded for the application of the verifiability rule must be able to be seen as belonging to a physico-material external world. These reasons must justify not only the externality of an inferentially reached grounding fact constituted by arrangements of physical tropes and their combinations but also (more directly) the externality of its aspectual sub-facts as partial arrangements

of physical tropes and their combinations (Ch. IV, sec. 25). In other words: the kind of commonsense direct realism defended in the last sections, though intuitively correct, still does not seem to justify the way the magic trick is performed of *interpreting* (reading, understanding, projecting, displacing… it is hard to find the right word) our internal sensory psychological contents as external physico-material contents perceived by our senses in a way similar to the way we interpret a mirror image as a reflection of something external. That is, even by accepting that we perceive the external world directly by means of the veil of perception after answering the arguments of illusion and science, we still seem unable to explain what we do in order to rid ourselves of what is internally mental when speaking of the external entities that are objects of perception.

In my view, a more complete answer begins to appear when we press the question further. Suppose we ask: under what conditions are semantic-cognitive rules like the verifiability rule not only conceivable, but also effectively applicable to entities belonging to the so-called real external world? In other words: what are the conditions responsible for our awareness of the effective applicability of the rule to what we are allowed to call *mind-independent physico-material entities really existing in the external world?* In still other words: when do those phenomenal entities that we could otherwise be able to recognize as mere sensory contents (sensations, sense-data) become likely to be recognized as directly experienced external tropical contents beyond our actual or dispositional mental states? (– I primarily mean their recognition as perceived external properties, that is, as simple or complex external tropes… material objects as clusters of external tropes displaying compresence… and real factual contents as external tropical arrangements.)

My suggestion is that what makes semantic-cognitive rules effectively applicable to mind-independent third-person physico-material tropical entities in the external world *is the satisfaction of suitable conditions of external reality in the absence of any verified skeptical scenario.* I hold that the adequate satisfaction of these conditions is ultimately responsible for the 'magic' of using our internal phenomenal sensory data as a way to reach external reality. That is, like the changes of our reading of mirror-images, our reading of what is phenomenally given in indexical thought-contents supported by sensory-impressions (internal criteria) can be changed into our reading of them as real factual contents belonging to the external world and constituted by what might be called physical external tropes (independent external criteria). More specifically, I wish to show that *by definition*, once the conditions for external reality are adequately and sufficiently satisfied, they constitute the proper, independent, externally

given *criterion* for the external reality of the contents of experience that fall under the scope of those rules. These conditions act somewhat like the conditions that, once considered, allow us to understand what we see in a mirror as the objects reflected and not as mere images of objects. (A skeptical scenario verified as a simulacrum of reality would be like a second mirror interposed between us and the object. The question is if the doubt whether there is not a second mirror makes any sense when there is no evidence for the existence of a second mirror.)

However, are there such conditions? In my view, these general conditions certainly exist, and their adequate satisfaction always constitutes what we implicitly assume in our attributions of external reality. The point was already touched on in the explanation of Mill's complementary conditions for external reality in Chapter IV. In fact, conditions for external reality were (within a diversity of metaphysical frameworks) already largely suggested by modern philosophers, beginning with Descartes and continuing with analytical philosophers, from G. E. Moore to J. R. Searle.

I can summarize the most fundamental conditions to warrant external reality proposed by modern philosophers, beginning with Locke. According to Locke, our opinions about physical objects are justified by the properties associated with our ideas of sensations, such as their involuntary character, order, coherent agreement reflecting law-governance, and interpersonal accessibility (1690, Book IV, Ch. 11). The immaterialist Berkeley concluded that the ideas constituting so-called external reality are very strong, distinct and independent of the will (1710, III). For Hume the impressions of a real thing are those that 'enter into the soul with the most force and violence' (1738, Book I, sec. 1). Kant held that conformity with laws (*Gesetzmäßigkeit*) is what defines the formal aspect of nature (1783, § 16). J. S. Mill, as we have already considered in some detail, said the external world consists not only in continuous or guaranteed or certified possibilities of sensations, but also in their independence of our will and their conformity with the regularities of nature, such as the causal laws of physics (1889, Ch. XI). According to Frege – already an analytic philosopher – the externally objective realm (his *erste Reich*) has as a criterion of objectivity its interpersonal accessibility and independence of will, while its reality has as a criterion its spatiotemporal location (1918b). A direct realist analytic philosopher of the early 20th century, G. E. Moore, summarized the main conditions of external reality in the following passage:

> The *real* is something independent of the mind that is verifiable by others, continuously connected with other things, and in this way has certain

> causes, effects and accompaniments with the highest degree of reality. (1953)

Such explanatory efforts have continued up to the present. To give an example, in a recent study John Searle pointed to some characteristics of the object or state of affairs really perceived, such as presentation (instead of representation), causation, non-detachability, indexicality (things are presented here and now), continuity and determinacy (Searle 2015: 60-70; See also Huemer 2001, Ch. 4).

Finally, a genetic account of our awareness of external reality proposed by Sigmund Freud (1911) could be mentioned. He suggested that we begin our lives under the governance of the *pleasure principle* (*Lustprinzip*), which seeks immediate gratification of desires and avoidance of pain. Since the external world does not grant us painless immediate gratification, we gradually learn the *reality principle* (*Realitätsprinzip*), according to which we need to act rationally toward the external world, postponing the immediate satisfaction of our desires in order to assure continuing lower levels of gratification accompanied by a foreseeably lower level of pain. For Freud, it is by means of this slow and difficult transition to the reality principle that we learn to distinguish an external material world with its own constraints.

It is true that when considered in isolation none of these conditions warrants that the contents of perceptual experience are externally real 'material' contents constructed from tropes. Indeed, criticizing Locke, Laurence BonJour correctly noted that none of the conditions of reality given by Locke is sufficient to warrant the external reality of anything (BonJour 2002: 130-135).

Examples easily confirm BonJour's objection: A mere content of sensation can have the highest degree of intensity and determinacy and yet be hallucinatory, as may occur in some rare cases. A perfectly realistic dream could be in strict conformity with all the expected regularities of our physical and social world. Although many mental acts are dependent on our will, dreams, obsessive thoughts, along with most feelings, are typically independent of our will. Even interpersonal agreement about states of affairs can occur without their real existence, as in the case of a dream in which we find other persons who agree or disagree with our experiences, or in the rare case of a collective hallucination (suppose that several people with similar beliefs take a hallucinogenic drug and, motivated by suggestion, share similar pseudo-perceptions…). Finally, external occurrences can possibly be directly dependent on our will (as in the case where someone has a brain-reader connected with his motor

cortex, enabling him to move objects in the outside world using his mind alone).

28. Answer: a definitional criterion of external reality

Notwithstanding, I think there is a way to surmount the problem identified by BonJour. The mere conditions of externality can be transformed into a *definitional criterion* for the existence or reality of the external world outside us, that is, into a *sufficient* condition for the ascription of external reality in the typical realist sense of the word. This suitable definitional criterion consists simply *in the demand that the most relevant of these conditions should be satisfied* together, *in accordance with conventional peculiarities of the expected kind of entity* (property, object, fact…). Hereby we find a decisively subsumed criterion that, once given, allows the perceptual contents that satisfy a semantic-cognitive rule to be projected onto the physical world outside us, as externally existing tropes or constructions made up of them, which are by definition external, material and free of any psychological element, within a complete system of external entities also able to satisfy the same kind of rule in a similar way.

We can better establish this point by proposing that the entities that can be seen as externally real are those that suitably satisfy *all* the main conditions of reality. And when these conditions are put together in such a way that we could in the proper sense of the word speak of them as constituting *a definitional criterion of external reality*, they work as what could be called *axioms of externality*. Here is how this view can be explicated concerning perceptible entities surrounding us:

> DEFINITIONAL CRITERION OF EXTERNAL REALITY:
> In order for some *content of belief* to satisfy an e-thought-content-rule as *an externally real entity* (as a factual content minimally constituted by an object and a property) belonging to the physico-material world, it must satisfy the basic axioms of externality. In the standard case, the axioms that the entity must satisfy in order to be considered externally real must be the following:
>
> (i) The entity must be able to be given to the senses of cognitive subjects *in its most intense degree and detail*. In many cases, it must also be *co-sensorially* given in the most intense degree.
> (ii) The entity must (usually) be *independent of the will*.

(iii) While existing, the entity must always be an *object of perceptual experience and possible interpersonal agreement insofar as the adequate conditions are given* (namely, it must be the subject of what Mill called 'continuous, guaranteed or certified possibilities of sensations') – here we could also speak of, if not *actual,* at least *interpersonally possible perceptual experience*.

(iv) The entity *must obey the laws of nature*, displaying expected regularities within a larger context (one would not be restricted to physical laws; biological, psychological and even social regularities could be included…).

(v) The entity must be able to be seen as in some (even if indirect or extremely indirect) way *causally related* to any cognitive subject who applies the rule.

I am not sure that this list cannot be improved, and I am unable to order the axioms hierarchically. But I believe they are the most relevant ones. Moreover, my thesis is that once *all* these conditions are suitably satisfied, there is nothing in the world that can defeat the kind of external reality that we intend to attribute by means of them. Together they are *sufficient* for the attribution of external reality in the most proper sense of the word, which I call the *inherent sense*, although it can be contrasted and easily confused with what will later be called the *adherent* sense of external reality, applicable in skeptical scenarios.

I will use the following example to make it clear that taken together the axioms of externality constitute a sufficient condition of external reality. Right now I am working on my notebook computer. I am very sure that what I am experiencing this device is presented to me as a complex of mental images and sensations (as variable contents of sensations or 'sense-data'). But I am also well aware that this device is also given to me as a corresponding *physico-material object existing externally* (an interpersonally perceivable and independently existing combination of material or external tropes of solidity, volume, form and… inertial mass, displaying compresence). Now, how do I know that I effectively apply my notebook's identification rule in its proper context, so that I am entitled to say that it exists externally? What warrants my understanding of the perceptual content as that of a real physico-material object in the outside world to which I might definitely apply its identification rule? The answer is clear: the suitable satisfaction of the above listed axioms of externality in the application of my notebook's semantic-cognitive identification rule. That is:

(i) The device must be able to be given to my senses in the most intense degree and detail. It is in this way also co-sensorially given (I can see, touch and hear it).
(ii) The device must exist and be constituted independently of my will.
(iii) The device must be continuously able to be given to sensory-perception under suitable conditions (I have seen this notebook computer intermittently in my home for many months and some other persons have also seen it). And (because of similar past experiences of material objects) I am sure that other persons would agree that this notebook computer is here in front of me now if they were here to see it; so I can be sure that my experience is at least interpersonally possible.
(iv) The device must obey expected physical regularities (it functions as described in the instructions, sometimes I have to recharge or replace its battery, I can download and install new programs from the internet, etc.)
(v) The device must satisfy its identification rule in a causal way (very often I am causally interacting with my notebook computer, and I am aware of this).

The interesting point is that even if the whole world were just an incredible dream – including my body and my notebook computer – I would still be entitled to affirm that my notebook computer is *indeed very real*, that it exists in an external world (even if in the end only a fictional one) and behaves externally as a material object in the sense that it suitably meets the criterion of reality by satisfying the axioms of externality from (i) to (v), that is, it is fully real in the *inherent* sense of the word. Indeed, if a dream has all the features of reality, then it is real in the relevant sense of the word.

A second point is that the satisfaction of the axioms may be incomplete and more or less constrained by conventions. A rainbow does not completely satisfy axiom (i): Although it is seen in its greatest expected intensity, it cannot be touched or heard. But probably for the same reason we aren't inclined to say that a rainbow is the most real thing in the world. New technology for brain-computer interfaces (BCI) enables us to move objects with willpower alone, which shows that some external things are to a certain extent dependent on our will and do not satisfy axiom (ii). Anyway, it is not our will that sustains their existence – not yet. And the real form of a mountain is conventionalized as satisfying the axioms of externality when viewed from its base and at a certain distance, which is a contextually dependent addition to the axioms.

On the other hand, internal sensory-contents, even those of a hallucination, typically do not satisfy, or only barely satisfy, the criterion of inherent reality. Indeed, if they sufficiently satisfy all externality axioms, they satisfy the criterion of external reality and must be considered in a sense externally real. It must be so because the totality of the partial conditions constituting the axioms of externality simply form a *definitional criterion*, grammatically or logico-conceptually warranting that the object of perception – in this case, the notebook computer in front of me – can be said to be an externally real material object belonging to what I am allowed to call a material physical world around me, and not just something merely mental.

Another point is that, although taken together the axioms of externality are a sufficient condition for attributions of external reality, their satisfaction is *not a necessary condition*. For instance, a person can be under the influence of some drug or suffer from some perceptual deficiency… so that although she is indeed experiencing a state of affairs that is externally real, several of these conditions are not being satisfied for her (e.g., some people take drugs to escape the harsh *reality* of the external world).

Anyway, it seems clear that it is the satisfaction of the externality axioms from (i) to (v) that for conceptual reasons alone warrants to me that my notebook computer's identification rule is effectively applicable to a real material object in the external world. Together these axioms establish the criterion for the application of our usual concept of the inherent reality or existence of things belonging to the external material world: a pre-condition that must be satisfied for the effective applicability of semantic-cognitive rules to inherently real things belonging to the domain of external physico-material reality. Their satisfaction warrants to me that a physico-material object like my computer *is real*, that it *exists externally* in a very concrete sense of the word, as a compresent cluster of stable tropes of solidity, density, volume, form, colors… effectively satisfying its identification rule and being constituted in conformity with it – and it must be so as a derivative of the cognitive senses we give to our words.

At this point a Berkeleyan immaterialist can object that even if the semantic-cognitive rule, along with the sensory content that satisfies it, demonstrates itself to be effectively applicable and consistent with all the axioms of externality, these *sensory contents* still belong to a mental order, so that although we believe that we grasp the material world, we remain in the domain of idealism. The answer I can give is that the sensory content that satisfies all these conditions must be a perceptual content belonging

by definition to an external physico-material world, insofar as we leave out of consideration radical skeptical doubts. My personal computer, by satisfying the axioms of reality, must satisfy the physical laws, it must satisfy the conditions for a material object (inertial mass inclusive) within the context of a physico-material world, at least insofar as no skeptical scenario has been found. The *psycho*-phenomenal content of sense data is now read as a physico-phenomenal content of material properties and objects. (I think that this point is made easier to understand when we remember that science has unmistakably proven that the mental is also physico-material, though as such only internally accessible.)

Skeptics will certainly object to this conclusion. They will point out that they can imagine skeptical hypotheses like those of a brain in a vat, a Cartesian soul or a dreaming subject… who are continuously and systematically being misled about a whole world that perfectly satisfies all the usual axioms of externality without having the least bit of external reality or the expected kind of physico-materiality. But as will be made clear below, the concept of external existence or reality applied in a skeptical scenario has an adherent sense, which is very different from the inherent sense of reality analyzed until now and, as we will see, cannot be applied in the absence of skeptical scenarios.

Now we already know what to do in order to warrant perceptual content, which as I noted is physical, like the content of a drawer. First, it is important to remember that the sensory-experience of mere sensory content is usually also a cognitive experience. If I have the feeling that a stove is hot or that I am holding a tennis ball, if I seem to hear a thunderclap or become aware of my sensory experience of my personal computer, these are all indexical thought-contents with their own verifiability rules satisfied through internal sensory experiences or sense-data. The point to be emphasized is that these sensory data will be read as internal only insofar as they are not seen as satisfying the criterion of external reality constituted by the axioms of externality (I may simply be hallucinating my notebook), while they will be read as external when they do satisfy these axioms (I have worked with this notebook computer for a long time, others have confirmed this). In the first case, I am considering the merely psychological experience (A) of sensory-psychological contents. In the second case, the applicability of the axioms of externality to what is given to me as sensory-psychological phenomenal contents (A) is only a transitive necessary condition for something further, namely, the proper perceptual experience (B) of external perceptual contents understood as physico-material tropical arrangements.

Furthermore, in case (B) we might suppose there to be something external *unifying* the variety of aggregates of sensory experience which make the real, actual external entity (a property, an object, a fact) in the world (for instance, my personal computer). Indeed, it seems reasonable to think that this actual external entity should have a unifying structure that could be captured through the effective applicability of a rule's many diverse criterial ramifications.

A final point is that in perceptual experience, when sub-facts sufficiently satisfy the verifiability rule, and this rule is accepted as effectively applicable because it satisfies the axioms of externality, we have at least a necessary condition for accepting the aspectual match between some derived indexical thought-content and the corresponding external sub-factual content. However, this satisfaction of the inherent sense of external reality also indirectly applies to the grounding fact represented by the basal thought, whatever it is. Consequently, for its effective application to facts belonging to the outside world, the verifiability rule must be applied in a way that also satisfy the axioms of externality. Only in such ways can ambiguous sensory-perceptual contents be understood as not merely mental, but as belonging to the inherently real external physical world, as physico-material constituents of the sub-facts belonging to a grounding fact, and because of this also to the grounding fact itself.

29. Proving the external world

Before we consider expected objections suggested by the consideration of skeptical scenarios, it is important to note that the application of the axioms of externality *can be inductively extended to contents that can be experienced only indirectly or potentially or both*. Thus, calling '(A)' the genetically originated trivial case of *the external reality of perceived entities surrounding us*, like the case of the personal computer I am writing on now, we also have other cases like:

> (A*) All the things we cannot experience directly with the unaided senses but can experience *indirectly*, such as viruses, atoms, magnetic fields, gravitational fields (one can indirectly verify the existence of atoms using scanning tunneling microscopy, and one can indirectly verify the existence of electromagnetic forces by manipulating magnetized material). These things can be considered externally real because the complexes of causes and effects that are associated with them satisfy axioms (i) to (v). Consequently, using a well-known

> mechanism of semantic extension already suggested by Aristotle, we are also justified in attributing external reality to them.[29]

Another form of semantic extension is case (B) of the application of the concept of *external reality to entities beyond the reach of our actual spatiotemporal possibilities of experience.* This case (B) can easily be subdivided into three subcases:

(B1) *Past things*. Everything I know to satisfy the criterion of external reality because I remember having experienced it as satisfying the criterion, but that is not accessible to me now (like my grandfather's house, which I visited only in my early childhood or a former childhood friend).

(B2) *Testimonial things*. The great number of things that I know are able to satisfy the criterion of external reality by means of testimony or any reliable informative source (from the city of Angkor Wat to Napoleon's coronation or the extinction of the dinosaurs). I would also include as 'testimony' photos, videos, historical documents, archaeological remains, etc.

(B3) *Unknown things*. This is finally the case of my inductive belief that because I have always had new experiences of real external things in the past, the world is full of other real external things that I have never experienced but that are directly or indirectly able to satisfy the criterion of external reality – let us call this 'the openness of the world.'

Finally, this allows us to inductively prove the inherent reality of the external world, since what we understand by *our whole world* is nothing more than the sum of all the entities that we reasonably believe to satisfy (A), (A*), (B) as (B1), (B1*), (B2), (B2*) and (B3) (B3*), understanding B1*, B2* and B3* as things indirectly experienced or able to be experienced in the corresponding domains.[30] In this way, we use the inherent criterion of reality in its extended forms to prove the existence or reality or actuality of the external world in the usual sense of the word.[31] It

---

[29] It is the same mechanism of semantic derivation that explains the archetypal truth-bearer considered in Chapter IV, sec. 30 of this book. See also Aristotle, *Metaphysics* 1003a, 33-37.

[30] For a more complete exposition of this point, see Costa 2014: 145-157.

[31] This is where Lewisanism – suggesting an infinite number of possible *real* worlds and accepting only our own world as real *and* actual – fails. We know that our existent, real, actual world, is distinct from merely possible non-real worlds,

is because all of us have unconsciously engaged in similar reasoning at some point in our childhood that we all believe that the external world self-evidently exists and that only philosophers and madmen can doubt its existence or reality.

These extensions also explain how we can make ordinary attributions of truth to statements based on adequation with inferentially derived statements of facts that aren't presently given to our senses. Consider as an example the judgment 'It is true that Mrs. Rose tried to poison Reverend David with arsenic,' symbolized as ⊢p, which is true by adequation with the inductively reached statement of fact symbolized by q. We know that q expresses a verifiability e-thought-content rule that can be read as representing an external factual content. But what entitles us to give the status of a fact to something that no person (with the exception of Mrs. Rose) ever observed? The answer is that we are inductively aware that this dynamic fact was real, satisfying the axioms of externality from (i) to (v) by the *indirect* means of the more direct satisfaction of the criteria of reality of the perceptual or perceptually based e-thought-contents r, s, t, and u, corresponding to their respective external facts. This entitles us to conclude that the verifiability rule of q would have been effectively applicable in its proper context if someone able to apply it were there, that is, that the fact-event of Mrs. Rose's attempt to murder her husband occurred as something inherently real.

30. Skeptical scenarios

Now, what about extreme skeptical scenarios or experiments with artificial reality? The challenge to our view is that in these cases the satisfaction of the definitional criterion of external existence considered above can (in principle) be *in part* or, it seems, *even totally emulated*. Thus, the brain in a vat (*pace* Putnam[32]) has experiences that seem as real to it as experiences

---

because we know that the verifiability rules of the facts of our world are (or could be) effectively applicable through the satisfaction of axioms of externality. In my view, D. K. Lewis' distinction between reality (inherent) and actuality (1986, Ch. 1) is a distinction without a difference that loses its teeth when we pay sufficient attention to the ways we can effectively attribute reality to a world.

[32] Hilary Putnam rejects the skeptical possibility that one could be a brain in a vat, hallucinating an unreal virtual reality produced by a supercomputer on the planet Omega or simply by chance (1981, Ch. 1). However, his objection is controversial, to say the least. According to Putnam's externalist point of view, if I am a brain in a vat, in order to have thoughts like those of brain, vat, water, etc., I need to be in *causal* contact with these things; hence, once I have these thoughts, I cannot be a

we have in our actual world, though it is on the very different planet Omega, interacting only with the program of a supercomputer... However, curiously enough, if this were the case and, for instance, the brain were removed from the vat and implanted in a living organism, so that it could experience the world of the planet Omega as it really is, *his past normal attributions of reality would in an important sense not be denied*. The same would be the case if someone comes to know he has been the object of a flawlessly executed virtual reality experiment. That is, happenings belonging to the life of the brain in a vat were very realistic illusions indeed, since the axioms (i) to (v) were all satisfied, even if everyone should agree that this solid reality was in a sense unreal, since it was a sub-product of the present world, here treated as if it were the 'ultimately real' world. Now, it seems that the world presented to the brain in a vat could be simultaneously real and unreal, which would be contradictory.

We can solve this dilemma by simply accepting that there are two different senses of external reality, which should not be confused:

(a) the *inherent* sense of external reality
(b) the *adherent* sense of external reality

The *inherent* sense of external reality (a) is the foregoing, demanding the suitable satisfaction of externality axioms from (i) to (v). We are all very well acquainted with the inherent sense, since it is the sense of reality that we apply on a daily basis. The brain in a vat (or the dreamer of a totally realistic dream) also experiences the criterion of inherent reality as being perfectly satisfied, and it is in this sense that the brain in a vat is right when it thinks that the experiences of the world given to it are perfectly real: they are still *real in the usual inherent sense*.

Nonetheless, the external world experienced by the brain in a vat before its liberation was *unreal in the adherent sense*, the sense (b) of external reality. The adherent sense is reserved for skeptical scenarios and

---

brain in a vat. The problem with Putnam's argument, as some have noted, is that it ignores the *flexibility* of language. Unless you are a staunch externalist on meaning, there is no good reason to believe that electrical patterns in the brain cannot misleadingly appear to us as brains, vats, water, etc. They could be falsely represented and intended as such, assuming that various *outside factors* (like the supercomputer on the planet Omega or anything belonging to a comparatively real external world) could systematically produce these patterns. Anyway, if you still believe in Putnam's argument you can choose another skeptical hypothesis like that of a realistic dream or appeal to 'recently envatted brains.' (See DeRose & Warfield, eds., 1999, Preface)

circumstances of virtual reality. It forms a different sense of 'external reality,' because the criterial conditions for satisfying or not satisfying the adherent sense of reality are very different from the criterial conditions for satisfying the inherent sense. The criteria for the adherent sense are more properly *coherential*. We would be able to reject the adherent reality of something experienced, based on the fact that we now know (or always knew) that we have been subjects of an experiment in virtual reality, since the coherence of that experience with actual and past surrounding circumstances is lacking.

Examples make the point clear. Consider an experiment with virtual reality in which we use special virtual-reality digital gloves that give us a sensation of touching holographic images of objects. You see the holographic image of a cup of tea, you touch the cup, you feel it, others can see it, but when you try to firmly grip it, your fingers go through the cup. Here to some extent the conditions of *inherent* reality are satisfied, though this does not suffice to endow the cup with reality. Moreover, we know from the start that the criteria of *adherent* reality are not being satisfied, since we are aware that the virtual reality is a counterfeit one made from material of our own real external world. We can even admit that the holographic image has some limited degree of inherent reality acquired by the satisfaction of some conditions of external reality, but surely no adherent reality and the reason for this last conclusion is that this evaluation fits much better with our more complete informational background.

Now, it is of utmost importance to see that when effectively applied in experience the concept of adherent reality is *relative*. Relative concepts are used only comparatively. For instance, consider the attribution of size in the sentence 'A small baby elephant is large relative to a mouse.' (Copi) Thus, we cannot effectively attribute or disattribute adherent reality independently of a given basis of comparison. Because of this, the idea that we can in a justified way know that the *ultimate* or *absolute* adherent reality of things is an empty one because it is devoid of criteria; we do not have and do not need to have a verifiability procedure assuring us that we and our actual world cannot be victims of a skeptical scenario. Because the concept of adherent reality is relative, we cannot prove the adherent reality of our world and we cannot disprove it. Consequently, to ask about the adherent reality or unreality of our world outside of any skeptical scenario is senseless. It is an illusion of reason invented by philosophers. We can ask only about its inherent reality and get a positive answer. That is all.

To further exemplify the relative or comparative character of the application of the concept of adherent reality, suppose now that you are a

brain in a vat. Let us suppose that you fall asleep one night, and when you wake up you find yourself in completely different surroundings with a new, unfamiliar bodily form. You see around you creatures that look quite strange and alien, and what is worse, you also look like them. They claim that you are now on the Planet Omega and tell you that your brain was removed from the vat and implanted in the head of a creature belonging to their species. The creatures give you coherential reasons to think that the world you are now living in is adherently real, *compared* with the world where you lived in the past, even if both are equally inherently real (they can show you the vat and the supercomputer. They tell you that the reason for the experiment is a pedagogical intention to increase the mental diversity on Planet Omega. They acquaint you with their wonderful new world, inhabited by the most fascinating creatures…). In the end (if you don't go insane) you may come to believe they are right, since this is the best way to give coherence to the relation between your present experiences and your memories. But it is important to remember that the application of the concept of adherent reality is here only comparative, since outside of the relationship to a skeptical scenario you cannot have any workable criterion to judge whether or not the present world is the ultimately real world. This impossibility is shown by the fact that even in a radical skeptical scenario where you have such a criterion it may be that you have been deceived again. Perhaps your brain was only moved to another vat, where the program 'Awaking on the planet Omega' is running, so that you are deceived again and will only be able to gain a new relative awareness of it if you are awakened once more…

On the other hand, the conditions of inherent reality are or have been equally well satisfied in any of these worlds, and in this sense they are all sufficiently real worlds. Thus, the earth-world was adherently unreal relative to the present Omega-world, while the present Omega-world is adherently real relative to the earth-world, even though both worlds are inherently real, and both worlds can turn out to be adherently unreal relative to a third adherently and inherently real world within a further skeptical scenario (e.g., being awakened from the program 'Awaking on Planet Omega') and so on.

These remarks are already sufficient to allow us to answer the radical skeptic, since it seems clear that for lack of semantic discernment the skeptic, as much as the anti-skeptic, confuses inherent attributions/disattributions of reality with relative adherent ones, producing *equivocal* arguments. According to the *modus tollens* skeptical argument for ignorance, because I cannot be absolutely sure that I am *really* not a brain in a vat, I cannot be sure that I have two *real* hands… However, here

the skeptic makes a mistake, since the concept of *reality* (usually implicit in the argument) should occur first in the adherent sense and then in the inherent sense. The anti-skeptic is victimized by the opposite confusion in his *modus ponens* argument for knowledge, according to which because I know that I have two *real* hands, I can know that the world where I now am is the *real* one and not a vat-world… since the (usually implicit) concept of reality here occurs first in the inherent sense and then should occur in the adherent sense.[33]

It is also important to note that the inherent reality of the external world experienced by the brain in a vat could not be one of a physico-material world obeying the laws of physics as we understand them! Indeed, being aware of a skeptical context, someone would be able to agree with the *lack of adherent reality* of the vat-world – made up only of electronic patterns in the supercomputer configured by a computer program – or the adherent unreality of the content of the world as a dream – constituted only by neuronal activity and not real material things surrounding the person who dreams. In a skeptical scenario, the attribution of adherent reality comes to the fore and makes sense, since there are reasons to make a comparison. But normally there is no reason. As we have already noted, this is why it is normally senseless to pose radical skeptical or anti-skeptical questions without offering a skeptical scenario, and this is why it is senseless to doubt or affirm that our world is a dream in the ultimate adherent sense. That is:

> The question 'Is our external world the ultimate, absolutely (adherently) real one?' is empty. It is a senseless transgression of the limits of meaningful language because it is an attempt to treat the relative concept of adherent reality as if it were an absolute (non-relative) concept.

Nonetheless, a question arises here: why are we so naturally disposed to accept the external world as not only inherently real, but also as the authentic physico-material world filled with the material objects we see around us, that is, as well as an adherently real world? Why is the assumption of the physical materiality of the external world part of our common knowledge? The answer is that people who ask this question

---

[33] Calling p any trivial proposition about the external world, s a person, h a skeptical hypothesis, and K the knowledge operator, the *modus tollens* skeptical argument has the form 1. $\sim sKh$, 2. $sKp \rightarrow sKh$, 3. Hence $\sim sKp$. The *modus ponens* anti-skeptical argument has the form 1. sKp, 2. $sKp \rightarrow sK\sim h$, 3. $sK\sim h$. (*Cf.* Costa 2014, Ch. 6.)

have not differentiated between the inherent and adherent senses of external reality. Because of this, they perceive that we can prove that the external world is real in the inherent sense, but they believe we are in this way also affirming that the external world is ultimately real in the adherent sense of reality. But they feel there is something excessive in this affirmation, which leads them to treat skeptical riddles as if they were more than mere semantic pseudo-problems. The question of adherent reality only arises because we are able to comparatively imagine skeptical scenarios in which the question would make sense. Inherent reality is all that we know on earth and all we need to know.[34]

The main point of this section was to reaffirm that adequate satisfaction of the axioms of externality is what essentially performs the sleight of hand of allowing us to interpret phenomenally given sensory contents as belonging to external physico-material entities independent of us, which by definition aren't mental or psychological. In this way, idealism is ruled out insofar as we find no evidence of a skeptical scenario providing us with relative criteria to pose the question of whether our world does or does not adherently exist and leading us to reject its physico-material reality. Once we feel ourselves free not only to interpret phenomenal contents as mind-independent, third-personally accessible, but also as obeying the real laws of nature, and therefore as being physico-material in all their aspects, we have no meaningful reason to pose the question of whether or not our world has adherent reality, simply because we lack verificational resources to answer that question, and a question without a possible answer is a question without meaning. Aside from skeptical scenarios, the satisfaction of the criterion of inherent reality by our phenomenal content is all that is needed to support the kind of displacement that puts content within what is called the non-mental external physico-material world.

31. Verification and intentionality: Husserl

At this point, it can be helpful to recall some of Edmund Husserl's views on truth in his *Sixth Logical Investigation*. I believe that he offers there his deepest insight, even if his insistent attempts to develop it might have

[34] One example of this kind of confusion is offered by Rudolf Carnap's conclusion that his external question about the existence of the external 'thing-world' in its totality must be answered by means of an irrational decision to accept the system, a pragmatic *fiat* (1947). For us, either this is an inherent question to be answered affirmatively or it is an adherent meaningless pseudo-question.

entangled him in a speculative maze. As we saw, Frege spoke of senses as meanings and thoughts, understanding them as abstract entities. The work of Wittgenstein, Michael Dummett, Ernst Tugendhat and others leads us instead to the suggestion that what Frege identified as senses or meanings are in fact semantic-cognitive rules or adequate associations of these rules considered in a *particularist* way as coming into being only through their effective or only merely rehearsed application. These rules can be applied either effectively (to the real world) or at least to some extent only imaginatively (as a possibility) if they do not remain mere psychological dispositions. Against this, Husserl spoke of *intentional acts* as ephemeral instantiations of meanings, supporting the Platonist view that meanings in themselves should remain abstract entities, as Frege and others have also held.

Nevertheless, it is important to see that Frege, Wittgenstein, and Husserl were all struggling with the very same issue, although using different strategies and starting from different perspectives and assumptions. As we saw, Fregean senses must be semantic-cognitive rules or associations of such rules. But similar reasoning should be applicable to Husserl's intentional acts: they should unavoidably include – in accordance with our view of semantics as always psychologically embodied – cognitive instantiations of semantic rules or associations of rules, which can be expressed in a cognitivist (psychological) and/or in a semanticist (logico-linguistic) fashion. As you might remember, in our analysis of adequation we considered an intention with a mind-to-world direction (responsibility) of fit added to its proper structural isomorphism as constitutive of a verifiability rule, which seems to a large extent a good way to understand Husserl's view of intentional acts.

In what follows, I will first present a short summary of Husserl's theory of intentionality in its relation to his adequation theory of truth. Then I will try to translate his main insights into my own conceptual framework.

As already noted, according to Husserl's view, the meaning (sense) of a linguistic expression is an *ideal*, an abstract (Platonic) object, as it was for Frege and others. However, for him the meaning of an expression can be instantiated by two fundamental kinds of ephemeral intentional acts:

(a) A *meaning-conferring intentional act* (*bedeutungsverleihende Akt* or *Bedeutungsintention*), which relates to an ideal object, abstracting its application to reality and disregarding truth-value (for example, I think that my sunglasses could be in the drawer);

(b) A *meaning-fulfilling intentional act* (*bedeutungserfüllende Akt*), which relates itself to the object actually given (for example, while looking for my sunglasses I open a drawer, where I find them).

In case (b) the object of the act is not only intended, it is also given to us 'in person,' even if always in perspectival ways, by means of distinct intuitions that can successively reinforce one another. Finally, there is a third act, an act (c) of *synthesis*, through which we make ourselves aware that the object intended in the meaning-giving intentional act is *the same* as the object intended as actually given in the meaning-fulfilling intentional act. For Husserl, with this last act we achieve awareness of truth and knowledge. Consequently, according to him, truth is correspondence because it is the identity of the object intended by the meaning-conferring act and the object intended by the meaning-fulfilling act. As he writes, truth is 'the complete agreement of what is intended with what is given as such.' (1980 vol. II/2, VI sec. 38) Knowing that there can be an unlimited variety of perspectival acts of fulfillment, which can be added to one another in order to warrant our knowledge of the object by giving the experience increasing evidential value, he also writes:

> When a presentative intention finds its ultimate fulfillment, the genuine *adaequatio rei et intellectus* is realized. The object is really presented as intended. So is the idea of all signitive fulfillment. The intellect is the intention of thought, the intention of meaning. Correspondence is realized when the intended object in the strict sense is given to us as it is thought. (1980 II/2 VI, sec. 37)

This 'correspondence' as the identity between the 'objects' of two intentions seems to me to be Husserl's chief insight on the nature of truth, since the process he describes is clearly at the origin of the pragmatics of adequation, as developed in the present chapter.[35]

Now, we can read meaning-conferring and meaning-fulfilling intentional acts as involving the instantiations of two verifiability rules. What Husserl identifies as the meaning-conferring intentional act can be approximated to the intention related to the verifiability rule that isn't effectively applied, but only taken into consideration – conceived as applicable. In other words, we see that it is possible for this rule to be definitely satisfied or applied, because we know by means of rehearsal that we can to a greater

[35] Husserl in fact distinguished four different concepts of truth. However, the question of their justification is controversial. (Husserl 1980, II, VI, sec. 39; *Cf.* Tugendhat 1970: 91 f.)

or lesser extent imaginatively apply it, as in the case of *?p*. On the other hand, what Husserl identifies as the meaning-fulfilling intentional act can be approximated to the intention related to a verifiability rule in its effective satisfaction or application within some actually given context. In the case where it is expressed by an assertoric sentence, this verifiability rule is a semantic-cognitive rule that can be said to be true or false in the sense that it can be shown to be effectively applicable or not. In the case in which we effectively apply a verifiability rule of the kind that can be expressed by an assertoric sentence, we are considering the act of *synthesis* by means of which the verifiability rule *?p,* due to its identity of content with *q,* is considered effectively applicable in its proper context, which also confers truth on *p* (⊢ *p*), making us aware of an actual fact that satisfies it.

32. Solving two Husserlian Problems

Now, comparing the kind of empiricist approach defended here with Husserl's theory of truth, we see that we are able to overcome two main drawbacks pointed out by his critics.

The first and more serious one is that working only with intentional-phenomenal material, Husserl was unable to explain the linkage of the object 'in person' with the object in itself, since this would require him to go beyond the phenomena. As Günter Patzig concluded:

> …the daring bridge called evidence intended to connect the judgment with the fact had the drawback, rather unfortunate in a bridge that it ended on the same side of the river from which it began. (1977: 194)

Our understanding of adequation offers us a non-idealist way to overcome this limitation. As already noted, the e-thought-rule expressed by *?p* can be approximated with what Husserl calls a meaning-conferring intention. And the e-thought-rule expressed by *!o* can be approximated to what Husserl called the meaning-fulfilling intentional act. Finally, the awareness of the qualitative identity of content represented by '*p* = *o*,' which brings us to the conclusion ⊢*p*, can be approximated to Husserl's synthesis by means of which we reach truth by seeing that the objects of the two acts are the same.

However, in doing this we do not need to follow Husserl in assuming some kind of idealism, because according to our analysis existence is the effective applicability of a conceptual rule, while the object of its application should only be conceived as what satisfies the criteria that could be generated by the rule, and its 'having existence or reality' is only

its potentiality of having its conceptual rule effectively applied to it. The same holds for the verifiability rule; this rule demands for its effective application the satisfaction of criterial configurations by isomorphically matching criterial configurations of the factual content belonging to the external world as it presents itself to us. These external criterial tropical configurations, in contrast, are manifestations of the empirical fact and are here not interpreted internally as psychological configurations of *sensory impressions*, but externally, as *real aspects* of external facts (that is, as tropes and constructions from tropes), insofar as they suitably satisfy the definitional criterion of external reality in its inherent sense. These are at least external aspects of what Husserl called the 'object in person,' but in our case, even if being sub-factual contents, they are externally real non-mental physico-material entities *by definition*. We can say that a fact is externally real because:

(1) This fact has the second-order dispositional property of having its first-order verifiability rule effectively applicable to it, even if this rule was never conceived or applied by any cognitive being.
(2) Insofar as the effective applicability of the verifiability rule to the external fact implies the satisfaction of the inherent criterion of reality that defines what is externally real in the most natural sense of the word (maximal intensity, independence of the will, interpersonal access, conforming to expected regularities, possible causal interaction…).

This fact will rightly be called a physico-material external fact, insofar as there is no skeptical scenario in view, for in the absence of a skeptical scenario there is no sense in questioning whether this external fact is not just inherently real, but also adherently real. It would be senseless, simply because the concept of adherent reality is a relative one, and the attempt to apply it in the absence of a skeptical scenario would be an attempt to transform it into an absolute concept – which can easily happen when the philosopher hasn't yet learned to distinguish inherent from adherent reality.

The second objection to Husserl's view is that the object is never given to us in its entirety. Since what we experience is always part of the object, it can never really be given to us 'in person.' Husserl saw this problem and

suggested that the object could still be seen as a pure or empty *X* of ideal nature (1976, sec. 52).[36]

Here I partially agree with him. Also, in the proposed view, it was assumed that neither the object nor the fact are perceptually given to us in their entirety, with the consequence that we can never be absolutely sure that what is given to our experience is the real object or fact. However, we can infer that the object or fact is given with enough probability, with practical certainty, assuming or postulating as warranted the evidence provided by the factually interpreted *!o* and, consequently, the corresponding truth of *p* in the context of an adequate linguistic practice, assuming that all other things remain the same. As we saw, we can infer that we have seen a dolphin and not just a rubber dolphin gliding over the water, and we can postulate what is given to our experience as indisputable evidence, insofar as we assume that the context of the expected observational practice is undefeated by unaccessed information.

Anyway, I agree that the compresent clusters of tropes that constitute the objects, as much as the linked property-tropes and the resulting facts, are in themselves inexhaustible. And this means that we can never be absolutely certain that any of our semantic-cognitive criterial rules is able to match such objects, properties or facts in order to warrant their existence in an unchallengeable sense. However, the fact that we cannot be absolutely sure of the external reality of what we have accepted as an external entity isn't sufficient to justify concluding that this entity must be something belonging to a purely mental realm. Indeed, there is a world of difference between the internal mentally-phenomenal (the 'phenomenological' of philosophy) and the external materially-phenomenal (the 'natural phenomena' of empirical science), which is conceptually warranted by satisfaction of the axioms of externality, insofar as there is no relativizing skeptical scenario in sight.

33. Truth and factual existence again

Now we return to the problem posed at the beginning of this chapter. There we asked whether the existence of a fact isn't the same thing as its

[36] Peter Simons summarized Husserl's view of intentional objects as follows: 'In particular, each noema has a kernel or nucleus which consists of three elements: a substratum, a set of qualitative moments, and modes of fulfillment of these qualities. What he calls a pure or empty *X* is the subject of predicates that are intended in the nucleus and which are more or less intuitively fulfilled. …this *X* is not a further concrete constituent in the noema; it is an abstract form occurring in it.' (Simons 1995: 127)

truth, since truth is also a property of a verifiability e-thought-content rule of being effectively applicable to a fact, which we have also understood as a correspondence with a fact, as was expressed by the formal identities (3) *T'p'* = *C'p'* = *V'p'* and (4) *'p'T*'q'* = *'p'C*'q'* = *'p'V*'q'*.

Nonetheless, we have also seen how to recognize here a false dilemma. 'Truth' in its proper sense of correspondence, as an e-thought-truth (propositional truth), can exist only as the result of the direct or indirect awareness of the effective applicability of a verifiability rule by at least one cognitive being, as we have clearly shown in our many examples of the dynamic processes that lead us to regard a verifiability e-thought rule as true or false. This amounts to the same thing as to say that the e-thought-rule represents its corresponding fact. Consequently, the variables *V* and *V** should be understood as abbreviations of such verifiability procedures. On the other hand, what we call a fact-truth, the existence of a fact, requires the effective applicability of its verifiability rule independently of our awareness of it, and thus even independently of the very instantiation of the rule in some mind (Ch. IV, sec. 34-35).

As we also saw, this means that real or true facts do not require the existence of epistemic subjects, existing without requiring anything beyond the dispositional property of being the object of application of *possible* verifiability rules, while e-thought-rules cannot be true without consisting of verifiability rules that are effectively applicable because at least one epistemic subject exists in the awareness that they were or could at some point be effectively applied to the corresponding facts. Because of this, 'truth' is an epistemic term, while 'existence' is an ontological term. The ontological (fact-truth) exists independently of the epistemic, while the epistemic (thought-truth) requires the ontological (fact-truth) in order to exist, necessitating for this at least one epistemic subject as a thought-bearer. This is why, despite similarities, we attribute truth to e-thought-content-rules and existence in the sense of reality to the facts that can be represented by them, while we do not attribute reality to e-thought-rules in order to replace truth.

The distinction considered here helps us to better understand the difference between the truth of a thought-content (thought-truth) and the existence of a fact (fact-truth) in the verification procedure. Consider the identity of contents verified in $p = \underline{q}$. The existence or reality of the fact is assumed by $\underline{q}$ (representing a fact-truth), and the truth of the thought-content is expressed by $\vdash p$ (expressing only a thought-truth). Even if *p* and *q* have qualitatively identical semantic contents (i.e., express identical verifiability rules) in the case of a true statement, the fact that they are

differently identified on the symbolic level points to the already indicated more substantial difference.

The role of the thought-truh can be grasped in a more detailed way if we again consider the truth-making procedure described in the case of Mrs. Rose's unfortunate husband:

$$!\underline{r} \quad \sim\!> ?p, \{!\underline{r} \,\&\, !\underline{s} \,\&\, !\underline{t} \,\&\, !\underline{u}\} \sim\!> !\underline{q}, p = \underline{q} \,/\, \vdash\!p$$

This whole action-schema presents a verification procedure constitutive of the e-thought-content-rule of p endowed with truth. That is, the verifiability e-thought rules expressed by r, s, t, u… are at least partial constituents of the thought-content of p – of its whole cognitive meaning. And I say 'partial constituents' because there are certainly many other ways to verify p, many other possible ramifications of the verifiability procedure. Moreover, r, s, t, u also have their own separate verification procedures constitutive of their own e-thought-content-rules besides the indispensable central e-thought-content-rule of p, which would be the direct verification of Mrs. Rose's attempts to poison her husband, which isn't available to us. Meaning comes to be an extended, gradually fading rule-complex, but since the above procedure is dependent on the direct verifiability of Mrs. Rose's attempts, the conceivability of the last one comes to be an indispensable meaning-condition. Finally, if Mrs. Rose confesses her attempts to murder her husband, we have added something very relevant for p's truth.

34. The rule's structural mirroring of the world

Let us recall that for J. S. Mill material substance was the 'permanent or warranted possibility of sensations' (Chap. IV, sec. 20). We have corrected this idea. Not the matter or substance, but the *existence* of the material object should be approximated to its permanent possibility of sensations, since permanence is always the same property, while objects can be endlessly diverse. Or, in our paraphrase, external existence is the effective applicability of the semantic-cognitive rule to entities of its proper domain or context, this effective applicability being measured by the assumed satisfaction of a criterion of inherent reality. This suggests a question: shouldn't for Mill matter or substance most properly be the multiple and variable configurations of 'sensations,' insofar as they are permanently accessible to our experience? Or, in our more qualified direct realist paraphrase, aren't the material entities (objects, their properties, the facts composed by them) constituted by the countless variable objective

configurations of external physical property-tropes able to suitably satisfy the axioms of externality necessarily required for the effective applicability of their semantic-cognitive rules in the external world?

The answer to this question seems to be: 'yes, but not only.' Indeed, dependent criterial configurations demand their isomorphic match with independent external criterial configurations enabling the application of semantic-cognitive rules, that is, mainly physico-material external quality-tropes and constructions from them (objects, facts) that are able to satisfy the rules, along with the expected satisfaction of the axioms of externality, since this allows us to classify such tropes and combinations of independent tropes as belonging to the external, material world.

This we already know. However, if it were only this, how could these multiple and diversified configurations of tropes that satisfy the criteria for the application of semantic-cognitive rules be conceived as belonging to *only one entity* (a complex property, a material object, a fact)? What is the glue that holds them together? How could they be *unified* instead of remaining inevitably dispersed? The plausible answer has been already suggested:

> What unifies all the aspects of an objective entity (property-tropes, individuals, facts) should be logically structured in a way that mirrors the logical structure of the semantic-cognitive rule.

Only in this inverted way would external structures be able to unify the multitude of external criteria. They are the totality of external criterial configurations, only a few of them being the configurations of tropes used to satisfy – that is, isomorphically match – dependent criterial configurations, though understood as belonging to the domain of the external world by satisfying the criterion of inherent reality.

In more detail: an objective external entity, be it (i) only a trope (complex or not, monadic or n-adic), be it (ii) a nuclear cluster of tropes displaying compresence and having the specific tropical properties constitutive of a material object, or be it (iii) any fact primarily conceived as a tropical arrangement (in the given case inevitably including (i) and (ii))… they should respectively mirror the *same logical structure* of the semantic-cognitive rules by means of which we ascribe predicative terms to (i), identify (ii) with nominal terms, and represent (iii) with statements.

This is why we can apply semantic-cognitive rules to a number of facets or aspects of an external entity and by these means identify the same entity as a whole; this is why a basal e-thought-rule can by means of its component rules be isomorphic with the elements of a grounding fact. This

is only possible because we assume that the perceived facets or aspects are associated with unperceived facets or aspects in ways that are *structurally similar* to those of the corresponding semantic-cognitive rules. In Chapter IV we used the rough metaphor of two identical trees that touch one another at the tips of their ramifications: on the one side, the dependent criterial configurations generated by the rule, on the other, the external ones – the structured configurations of material/external tropes (possibly complemented by mental/internal tropes, as in many complex physico-social states of affairs).

A trivial example can show the plausibility of the idea that a semantic-cognitive rule's logical structure should mirror the logical structure of the entity to which it applies, which on its side should be mirrored in the structure of the rule. Suppose that one day I start driving to the university, where I intend to hold a class. As I drive onto the freeway, I see that there is less traffic than usual. I begin to ask myself if today is a holiday.[37] I do not have with me any smartphone to check whether today is a holiday. However, some minutes later I arrive at the university where I find that it is closed, and a security guard tells me that today is a national holiday. Now, I have used ramifications of the verifiability rule to confirm the truth of (I) 'Today is a holiday.' This is confirmed by three facts: the symptom (a) that there is less traffic than usual on the freeway, the secondary criterion (b) that the university is closed; and the (less) secondary criterion (c) that when asked, a security guard informs me that it is, in fact, a national holiday. From the thought-content of (i), I derived ramifications of the verifiability rule which were the thought-contents of (a), (b) and (c). But on the other hand, I can say that from the corresponding institutional fact that it is a holiday, more completely stated as the grounding fact that today is a national holiday, which was declared to be one by Congress and was institutionalized as a law by publication in the official legal gazette... From this grounding fact (I*) inductively follow sub-facts that can be used as symptoms or secondary criteria, such as (a*) there are fewer cars than usual on the roads, (b*) the university is closed, and (c*) if one asks a security guard, he will certainly say that today is a holiday. That is: the same things that inferentially follow from statement (I) as its verifying criteria or symptoms also follow from the institutional grounding fact (I*)

[37] I do not consider the many ways I have to verify this hypothesis; but I know very well the implications of its falsification. One of them is that I will not hold any class today; another is that I am wasting my time going to the university. As we have seen (Ch. V, sec. 1), these inferences are more or less derived from my *awareness* of the meaning of the supposition that today is a holiday, though they do not belong to its meaning.

that today is a holiday, allowing a corresponding multiplicity of matches. And this makes it sufficiently clear that the ramified logical structure of the applied verifiability rule mirrors similarly ramified structural relations of sub-facts derived from the grounding fact that today has been declared a legal national holiday.

Nonetheless, it is also fundamental to perceive that usually our awareness of most of these mirrored structures is *merely putative*. The structure of objective reality is often more complex or is only approximately similar to that of our semantic-cognitive rules. Indeed, we usually assume that our semantic-cognitive rules are inevitably fallible, insofar as they are directed at the open world of experience. That is, we only *assume as probable* that the structures of the internal semantic-cognitive rules mirror the structures of their external unifying references, which can in principle be corrected or even refuted by new experiences, leading us to expansions or disavowals regarding the structures of the semantic-cognitive rules. This can be the case with ascription rules, identification rules, and verification rules, and is more explicitly shown by rules stated as laws of nature.

Summarizing: material objects, complex property-tropes and facts must have *proper unifying logical structures* that explain why the entities in question remain the same, even when experienced in different ways; and these structures are thought to be mirrored by the many variable structures of the semantic-cognitive rules that allow us to refer to them in unified ways.

35. Conclusion

The conclusion of this chapter can be extended to the whole book. It was an attempt to restore and unify some unjustly undervalued but intuitively fundamental ideas of the linguistic-analytic tradition. Once their acceptability is revealed, it is easier to see where they can be related to one another, building in this way what seems a plausible systematic overview. If the arguments presented here are essentially correct and well-grounded, then analytic philosophy of language should follow a different trajectory, and the right method to learn could well be that of ‘successive approximations’ (Haack 2016), instead of almost gratuitous counter-intuitive challenging – going further by correcting and detailing the rough sketches presented here.

APPENDIX TO CHAPTER VI

DISCOVERY OF WINE

Notre époque est une époque de misère sans art, c'est pitoyable. L'homme est nu, dépouillé de tout, même de sa foi en lui.
[Our time is a time of misery without art, it is pitiable. Man is naked, despoiled of all, even of his faith in himself.]
—Louis-Ferdinand Céline

The name of poet was almost forgotten; that of orator was usurped by the sophists. A cloud of critics, of compilers, of commentators, darkened the face of learning, and the decline of genius was soon followed by the corruption of taste.
—Edward Gibbon

We have first raised a dust and then complain we cannot see.
—Berkeley

Once one absurdity is accepted, the rest follows.
—Aristotle

There is a mythical story of the discovery of wine, told by the humorist Millôr Fernandes in his book *Fabulous Fables* (1963), which I would like to recall here. It goes as follows:

A traveler once needed to cross a desert. Since he loved grapes, though not grape pits and skin, it occurred to him that in his saddlebag he could bring with him not water but only the juice of crushed grapes. After a journey of three days, he noticed that the juice had turned yellow, tasted different and was releasing bubbles. After he had drunk this beverage, he noted that it made him feel much better than usual, so much so that he judged this to be the most enjoyable trip of his entire life! After his arrival, he told the news to his fellow travelers, who decided to follow his example, making long journeys across the desert with heavy saddlebags filled with the juice of crushed grapes, so that they could enjoy the same feeling of well-being. For a long time, this state of affairs continued unchanged, until one day a stubborn camel refused to commence a journey and for three days remained so to speak nailed to the same spot with the

grape juice on his back. To the surprise of the camel's owner, the juice also changed its color and taste and released bubbles. The news spread quickly. From the discovery of fermentation to the bottling and commercialization of wine, it was hereafter a very short step.

For me, this story illustrates the all too easy ways in which we can go astray in developing our philosophical conjectures. In a world plagued by a growing multiplication of philosophical views, many of them inevitably standing on deeply equivocal foundations, this difficulty can lead to dangerous disorientations that accumulate in unappealing forms of escapism like skepticism, relativism, irrationalism and most unexpected forms of sophistry[1]… not to speak of expansionist scientism. Under these circumstances, the effort to achieve some comprehensive picture – as was attempted here – could probably (though not certainly) furnish better guidance by suggesting conclusions that have a better chance to impose themselves by the cumulative force derived from the picture's internal coherence.

My point should not be carried too far: the above mentioned pitfalls have in one way or another always belonged to philosophy, since it has always encouraged hopes that have subsequently been exposed as highly questionable, even in the best of cases. In this regard, my hope concerning the stories told in this book is still the same: I *believe* I have approached the right comprehensive view of the cluster of conceptual structures centered in the notions of reference and cognitive meaning, finding in this way the path that could lead us to critical consensual truth. I see it as a multi-perspectival alternative approach benefiting from some of the best insights of the history of analytic and traditional philosophy. Something that should cut deeply into the inherited wisdom of contemporary philosophy of language, having the potential to liberate it from its main stalemates and to remap much of the field by bringing it back to its most proper epistemic center. In this way it might offer us renewed hope of approaching science in its liberalized sense as 'consensualizable public knowledge' (Ziman 1968),[2] sparing the reader many long, senseless journeys across the scorched desert sands of philosophically illusory arguments.

---

[1] See, for instance, Frankfurt 2005.

[2] This is an all-embracing intuitive definition of science as any knowledge already able, within the appropriate community of ideas, to achieve legitimate consensual truth concerning its results. This has been impossible for philosophy due to a lack of consensus regarding fundamental assumptions concerning methodology and starting points (Costa 2002, Ch. 2). But, if the views defended in the present book are substantially right, we now have a better chance to lift some issues of theoretical philosophy to a less speculative stage.

REFERENCES

Adorno, T. and M. Horkheimer (2010). *Die Dialetik der Aufklärung*. Frankfurt: Fischer Verlag.

Alston, P. W. (2001). 'A Realist Conception of Truth,' in M. P. Lynch (ed.), *The Nature of Truth: Classic and Contemporary Perspectives*. Cambridge, MA: MIT Press.

Appiah, K. A. (2002). *Thinking it Through: An Introduction to Contemporary Philosophy*. Oxford: Oxford University Press.

Aristotle (1984). *The Complete Works of Aristotle*, (ed.) J. Barnes, 2 vols. Princeton, NJ: Princeton University Press.

Armstrong, D. M. (2010). *Sketch for a Systematic Metaphysics*. Oxford: Clarendon Press.

—. (2004). *Truth and Truth-Makers*. Cambridge: Cambridge University Press.

—. (1999). *The Mind-Body Problem: An Opinionated Introduction*. Boulder: Westview Press.

—. (1997). *A World of States of Affairs*. Cambridge: Cambridge University Press.

—. (1981). 'What is Consciousness?' in *The Nature of Mind*. Ithaca, NY: Cornell University Press.

Audi, R. (1998). *Epistemology*. London: Routledge.

Augustine, St. (2008). *Confessions*. Oxford: Oxford University Press.

Austin, J. L. (1962). *How to do Things with Words*. Oxford: Oxford University Press.

—. (1962). *Sense and Sensibilia*, (ed.) G. J. Warnock. Oxford: Oxford University Press.

—. (1961). 'Unfair to Facts,' in *Philosophical Papers*. Oxford: Oxford University Press.

—. (1950). 'Truth.' *Proceedings of the Aristotelian Society*, 24, 111-128.

Ayer, A. J. (1992). *The Philosophy of A. J. Ayer*, (ed.) L. E. Hahn. La Salle, I.L.: Open Court.

—. (1973). *The Central Questions of Philosophy*. London: Weidenfeld & Nicholson.

—. (1972). 'One's Knowledge of Other Minds,' in *Philosophical Essays*. London: Macmillan & Saint Martin's Press.

—. (1963). 'Truth.' In *The Concept of Person and Other Essays.* London: Macmillan Press.
—. (1959). 'Verification and Experience,' in (ed.) A. J. Ayer, *Logical Positivism.* New York: The Free Press.
—. (1952). *Language, Truth, and Logic.* New York: Dover Publications.
Baars, B. (1997). *In the Theater of Consciousness: The Workspace of the Mind.* New York: Oxford University Press.
Baker, G. P. (1974). 'Criteria: a New Foundations for Semantics.' *Ratio* 16*,* 56-189.
Baker, G. P. & P. M. S. Hacker (1984). *Frege: Logical Excavations.* Oxford: Oxford University Press.
—. (1980). *Wittgenstein: Understanding and Meaning – An Analytical Commentary on the Philosophical Investigations.* Oxford: Oxford University Press.
Beaney, M. (ed.) (1997). *The Frege Reader.* Oxford: Blackwell.
Bell, D. (1990). *Husserl: The Arguments of the Philosophers.* London: Routledge.
Berkeley, G. (1979). *Three Dialogues between Hylas and Philonous* (1713), (ed.) R. M. Adams. Indianapolis, IN: Hackett.
—. (1975). *Of the Principles of Human Knowledge* (1710), in *Philosophical Works Including the Work on Vision*, (ed.) M. R. Ayers. London: Everyman.
—. (1975). *Philosophical Commentaries* (1707-8), *Philosophical Works Including the Work on Vision*, (ed.) M. R. Ayers. London: Everyman.
Black, M. (1954). *Problems of Analysis.* Ithaca, NY: Cornell University Press.
Blackburn, S. (1984). *Spreading the Word: Groundings in the Philosophy of Language.* Oxford: Clarendon Press.
Blanchard, B. (1939). *The Nature of Thought*, vol. 2. London: George Allen and Unwin.
Block, N. (1995). 'On a Confusion about a Function of Consciousness.' *Behavioural and Brain Sciences* 18, 227-47.
—. (1987). 'Functional Role and Truth-Condition.' *Proceedings of the Aristotelian Society*, 61, 167-181.
BonJour, L. (2002). *Epistemology: Classic Problems and Contemporary Responses.* Lanham, MD: Rowman & Littlefield.
—. (1998). *In Defence of Pure Reason.* Cambridge: Cambridge University Press.
—. (1985). *The Structure of Empirical Knowledge.* Cambridge, MA: Harvard University Press.

Bosanquet, B. (2015). *Truth and Coherence* (1911). Reprints from the collection of the University of Michigan Library.

Bradley, F. H. (1914). *Essays on Truth and Reality*. Oxford: Clarendon Press.

Brake, E., J. Millum (2016). 'Parenthood and Procreation.' *The Stanford Encyclopedia of Philosophy* (internet archive).

Brandon, R. B. (2000). *Articulating Reasons: An Introduction to Inferentialism.* Cambridge, MA: Harvard University Press.

Burge, T. (2005). 'Sinning against Frege,' in *Truths, Thoughts, Reason: Essays on Frege.* Oxford: Clarendon Press.

—. (1979). 'Individualism and the Mental.' In *Midwest Studies in Philosophy, vol. IV, Studies in Metaphysics*, (eds.) P. A. French, T. E. Uehling Jr., H. K. Wettstein, 73-121.

Campbell, K. (1990). *Abstract Particulars.* Oxford: Blackwell.

—. (1981). 'The Metaphysic of Abstract Particulars.' *Midwest Studies in Philosophy* 6, 477-486. Reprinted in S. Lawrence and C. McDonald (eds.) (1988). *Contemporary Readings in the Foundations of Metaphysics.* Oxford: Blackwell.

—. (1974). 'One Form of Scepticism about Induction,' in R. Swinburne, *The Justification of Induction.* Oxford: Oxford University Press.

Carnap, R. (1947). 'Empiricism, Semantics and Ontology,' in *Meaning and Necessity.* Chicago: University of Chicago Press.

—. (1937). *The Logical Syntax of Language*. London: Keagan Paul Trench & Trubner.

Chalmers, D. J. (2006). 'The Foundations of Two-Dimensional Semantics,' in M. Garcia-Carpintero, J. Macià, *Two-Dimensional Semantics.* Oxford: Clarendon Press.

Church, A. (1956), *Introduction to Mathematical Logic*. Princeton, NJ: Princeton University Press.

—. (1949). 'Review of Language Truth and Logic.' *Journal of Symbolic Logic*, 14, 52-53.

Copi, I. (1972). *Introduction to Logic.* New York: MacMillan.

Copleston, F. (1993). *A History of Philosophy.* New York: Doubleday.

Cornman, J. W. (1975). *Perception, Common Sense, and Science.* New Haven, CT: Yale University Press.

Costa, C. (2014). *Lines of Thought: Rethinking Philosophical Assumptions*. Newcastle upon Tyne: Cambridge Scholars Publishing.

—. (2011). *Paisagens conceituais: ensaios filosóficos.* Rio de Janeiro: Tempo Brasileiro.

—. (2002). *The Philosophical Inquiry: Towards a Global Account.* Lanham, MD: University Press of America.

—. (1997). 'Das Paradox der privaten Erfahrung,' in *Prima Philosophia*, 10, 433-448.
—. (1995). 'Fatos empíricos,' in *A Linguagem Factual.* Rio de Janeiro: Tempo Brasileiro.
—. (1990). *Wittgensteins Beitrag zu einer sprachphilosophischen Semantik.* Konstanz: Hartung-Gorre Verlag.
Crane, T. (1991). 'All the Difference in the World.' *Philosophical Quarterly* 41, 162, 1-26.
Dancy, J. (1985). *Introduction to Contemporary Epistemology.* Oxford: Blackwell.
Davidson, D. (1986). 'True to the Facts.' *The Journal of Philosophy*, 66, 748-64.
Dennett, D. (1993). *Consciousness Explained*. New York: Penguin Books.
—. (1982). 'Beyond Belief,' in A. Woodfield (ed.), *Thought and Object: Essays on Intentionality*. Oxford: Clarendon Press.
Descartes (1978). *Meditationes de Prima Philosophia – Méditations Métaphysiques* (1641). Paris: J. Vrin.
Devitt, M. & K. Sterelny (1999). *Language and Reality: An Introduction to the Philosophy of Language*. Oxford: Blackwell.
Devitt, M. (1997). *Realism and Truth.* Oxford: Blackwell.
—. (1990). 'Meanings just ain't in the Head,' in George Boolos (ed.), *Meanings and Method: Essays in Honor of Hilary Putnam.* Cambridge: Cambridge University Press, 1990.
—. 1981. *Designation.* New York: Columbia University Press.
Donnellan, K. (1974). 'Speaking of Nothing,' *The Philosophical Review*, 83, 3-32.
—. (1970). 'Proper Names and Identifying Descriptions.' *Synthesis* 21, 335-58. Republished as 'Names and Identifying Descriptions,' in Donald Davidson & Gilbert Harman (1972) (eds.): *Semantics of Natural Language*, Dordrecht: Reidel.
Duhem, P. (1906). *La Théorie physique: Son object et sa structure.* Paris: Chevalier & Riviére.
Dummett, M. (1993), 'What is a Theory of Meaning? (I)' and 'What is a Theory of Meaning? (II),' in *The Seas of Language.* Oxford: Clarendon Press.
—. (1981a). *The Interpretation of Frege's Philosophy*. London: Duckworth.
—. (1981b). *Frege: Philosophy of Language* (1973). London: Duckworth.
—. (1978). *Truth and Other Enigmas.* Cambridge, MA: Harvard University Press.

Dupré, J. (1993). *The Disorder of Things: Metaphysical Foundations of the Disunity of Science.* Cambridge, MA: Harvard University Press.
—. (1981). 'Natural Kinds and Biological Taxa.' *The Philosophical Review* 90, 66-91.
Elburne, P. (2011). *Meaning: A Slim Guide to Semantics.* Oxford: Oxford University Press.
Engels, P. (2002). *Truth.* Montreal: McGill-Queen's University Press.
Erwin, E., L. Kleiman, E. Zemach (1974). 'The Historical Theory of Reference.' *Australasian Journal of Philosophy*, 56, 50-56.
Evans, G. (1985). 'Semantic Theory and Tacit Knowledge,' in G. Evans, *Collected Papers.* Oxford: Oxford University Press.
—. (1982). *The Varieties of Reference*, ed. J. McDowell. Oxford: Oxford University Press.
—. (1973). 'The Causal Theory of Names,' *Aristotelian Society: Supplementary Volume* 47.
Fernandes, M. (1963). 'A descoberta do vinho' in *Fabulas fabulosas.* Rio de Janeiro: J. Álvaro editora.
Fitch, G. W. (2004). *Kripke*. Montreal & Kingston: McGill-Queen's University Press.
Fogelin, R. (1995). *Wittgenstein.* London: Routledge.
Forbes, G. (1990). 'The Indispensability of *Sinn*.' *Philosophical Review* 89, 535-63.
Foster, J. (2000). *The Nature of Perception*. Oxford: Oxford University Press.
Frankfurt, H. (2005). *On Bullshit.* Princeton, NJ: Princeton University Press.
Frege, G. (1918-1919a). 'Die Verneinung.' *Beitrage zur philosophische des deutschen Idealismus* 1, 143-157
—. (1918-1919b). 'Der Gedanke.' *Beitrage zur philosophie des deutschen Idealismus* 2, 58-77.
—. (1969). 'Einleitung in die Logik,' in H. Hermes, F. Kambartel, F. Kaulbach (eds.), *Nachgelassene Schriften.* Hamburg: Felix Meiner.
—. (1969). 'Ausführungen über Sinn und Bedeutung,' in H. Hermes, F. Kambartel, F. Kaulbach (eds.), *Nachgelassene Schriften.* Hamburg: Felix Meiner.
—. (1892). 'Über Sinn und Bedeutung.' *Zeitschrift für Philosophie und philosophische Kritik,* NF 100, 25-50.
—. (1892). 'Über Begriff und Gegenstand.' *Vierteljahreschrift für wissenschaftliche Philosophie* 16, 192-205.
—. (2008). 'Funktion und Begriff' (1891), in Günter Patzig (ed.), *Funktion, Begriff, Bedeutung.* Göttingen: Vandenhoek & Ruprecht.

—. (1988). *Die Grundlagen der Arithmetik: Eine Logish Mathematisch Untersuchung über den Begriff des Zahls* (1884). Hamburg: Felix Meiner.
—. (1977). *Begriffschrift und andere Aufsätze* (1879), ed. I. Angelelli. Darmstadt: Wissenschaftliche Buchgesellschaft.
Freud, S. (1927). *Die Zukunft einer Illusion.* Leipzig, Wienna, Zürich: Internationaler Psychoanalytischer Verlag.
—. (2009). *Das Ich und das Es: Metapsychologische Schriften* (1923). Frankfurt: Fischer Verlag.
—. (2014). 'Formulierungen über die zwei Prinzipien des psychischen Geschehens' (2011), in *Werke aus den Jahren 1909-1913.* Gesammelte Werke. Frankfurt: Fisher Verlag.
—. (1998). *Die Traumdeutung* (1900). Frankfurt: Fischer Verlag.
Glock, H-J. (2010). *Wittgenstein-Lexikon*. Darmstadt: Wissenschaftliche Buchgesellschaft.
Gower, B. (2006). 'A. J. Ayer: Language, Truth and Logic', in J. Shand (ed.) *Central Works of Philosophy*, vol. 4. Montreal & Kingston: McGill Queen's University Press.
Grice, P. (1989). *Studies in the Ways of Words.* Cambridge, MA: Harvard University Press.
Grice, H. P. & P. F. Strawson (1956). 'In Defense of a Dogma.' *The Philosophical Review* 65, 141-158.
Gulick, R. Van (2014). 'Consciousness.' *Stanford Encyclopedia of Philosophy* (internet archive).
Haack, S. (2017). *Scientism and its Discontents*. Rounded Globe.
—. (2016). *The Reintegration of Philosophy*. Springer Verlag.
—. (1978). *Philosophy of Logics.* Cambridge: Cambridge University Press.
Habermas, J. (1984). 'Wahrheitstheorien,' in *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns.* Frankfurt: Suhrkamp.
—. (1976), 'Was heißt Universal Pragmatik?' in K. O. Apel (ed.), *Sprach-pragmatik und Philosophie.* Frankfurt: Suhrkamp.
Hacker, P. M. S. (1990). *Wittgenstein: Meaning and Mind: An Analytical Commentary on the Philosophical Investigations*, vol. 3, part I. Oxford: Blackwell.
—. (1986). *Insight and Illusion.* Oxford: Oxford University Press.
Hallett, G. (1988). *Language and Truth*. Yale: Yale University Press.
Hamilton, A. (1988). 'Mill, Phenomenalism, and the Self,' in J. Skorupsky (ed.), *The Cambridge Companion to Mill.* Cambridge: Cambridge University Press.

Hare, H. H. (1981). *Moral Thinking, its Levels, Methods and Point.* Oxford: Clarendon Press.
Harman, G. (1999). '(Non-Solipsistic) Conceptual Role Semantics,' in G. Harman, *Reasoning, Meaning and Mind.* Oxford: Oxford University Press.
Hempel, C. G. (1967). *Philosophy of Natural Science.* Englewood Cliffs, NJ: Prentice-Hall.
—. (1959). 'The Empiricist Criterion of Meaning,' in A. J. Ayer (ed.), *Logical Positivism.* New York: Free Press. First published in the *Revue Internationale de Philosophie*, vol. 4 (1950).
—. (1935). 'On the Logical Positivists' Theory of Truth.' *Analysis* 2, 49-59.
Hilmy, S. S. (1987). *The Later Wittgenstein.* Oxford: Blackwell.
Hirsch, E. (1982). *The Concept of Identity*. New York: Oxford University Press.
Hughes, C. (2004). *Kripke: Names, Necessity, and Identity.* Oxford: Clarendon Press.
Huemer, M. (2001). *Skepticism and the Veil of Perception.* Lanham, MD: Rowman & Littlefield.
Hume, D. (2004) *An Inquiry Concerning Human Understanding* (1748). New York: Barnes & Noble.
—. (1975). *A Treatise of Human Nature* (1738). L.A. Selby-Bigge (ed.), 2a. ed. revised by P. H. Nidditch. Oxford: Clarendon Press.
Husserl, E. (1954). *Die Krisis der europäischen Wissenschaften und die transzendentale Phenomenologie: eine Einleitung in die phänomenologische Philosophie.* Den Haag: Martinus Nijhoff Publishers, Husserliana, vol. VI.
—. (1976). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie* (1913). Hamburg: Felix Meiner Verlag.
—. (1980). *Logischen Untersuchungen* (1901), 3 vols. Tübingen: Max Niemeyer Verlag.
Hylton, P. (2003). 'The Theory of Descriptions,' in N. Griffin (ed.), *The Cambridge Companion to Bertrand Russell.* Cambridge: Cambridge University Press.
Jackson, F. (2010). *Language, Names, and Information.* Oxford: Wiley-Blackwell.
Juhl, C., E. Loomis (2010). *Analyticity.* New York: Routledge.
Johnston, M. (1997). 'Manifest Kinds.' *The Journal of Philosophy*, 94, 564-83.
Kallestrup, J. (2012). *Semantic Externalism.* London: Routledge.
Kant, E. (1988). *Kritik der reinen Vernunft* (1787). Frankfurt: Suhrkamp.

—. 1988. *Prolegomena zu einer jeden künftigen Metaphysik, die als Wissenschaft wird auftreten können* (1783). Frankfurt: Suhrkamp.
Kaplan, D. (1989). 'Demonstratives,' in J. Almog, J. Perry & H. Wettstein (eds.), *Themes from Kaplan.* Oxford: Oxford University Press.
Katz, J. (1990). 'Has the Description Theory of Names been Refuted?' in G. Boolos (ed.), *Meaning and Method: Essays in Honour of Hilary Putnam.* Cambridge: Cambridge University Press.
Kenny, A. (1995). *Frege: An Introduction to the Founder of Modern Analytic Philosophy.* Oxford: Blackwell.
—. (1993). *Aquinas on Mind.* London: Routledge
—. (1973). *Wittgenstein.* Cambridge, MA: Harvard University Press.
Kirkham, R. (1995). *Theories of Truth: A Critical Introduction.* Cambridge, MA: MIT Press.
Kneale, W. & M. Kneale (1985). *The Development of Logic.* Oxford: Oxford University Press.
Kolakowsky, L. (2001). *Metaphysical Horror.* Chicago: University of Chicago Press.
Kornblith, H. (1980). 'Referring to Artifacts.' *Philosophical Review* 89, 109-114.
Kripke, S. (2013). *Reference and Existence.* Oxford: Oxford University Press.
—. (1982). *Wittgenstein on Rules and Private Language: An Elementary Exposition.* Cambridge, MA: Harvard University Press.
—. (1980). *Naming and Necessity* (1970). Cambridge, MA: Harvard University Press.
—. (1971). 'Identity and Necessity,' in M. K. Munitz (ed.), *Identity and Individuation.* New York: New York University Press.
Künne, W. (2013). *Conceptions of Truth.* Oxford: Oxford University Press.
Laporte, J. (2004). *Natural Kinds and Conceptual Change.* Cambridge: Cambridge University Press.
Larson, R., G. Segal (1995). *Knowledge of Meaning: An Introduction to Semantic Theory.* Cambridge, MA: MIT Press.
Lau J., M. Deutsch (2016). 'Externalism about Mental Content.' *Stanford Encyclopedia of Philosophy* (internet archive).
Leibniz (1921). *Nouveaux Essais sur L'entendement Humain* (1765). Paris: Flammarion.
Lemos, N. (2004). *Common Sense: A Contemporary Defence.* Cambridge: Cambridge University Press.
Lewis, D. K. (1986). *On the Plurality of the Worlds.* Oxford: Blackwell.

—. (1984). ‘Putnam’s Paradox.’ *Australasian Journal of Philosophy*, 62, 221-236.
—. (1969). *Convention: A Philosophical Study.* Cambridge, MA: Harvard University Press.
Loar, B. (1976). ‘The Semantics of Singular Terms.’ *Philosophical Studies*, 30, 353-377.
Locke, J. (1975). *Essay Concerning Human Understanding* (1690), (ed.) P. H. Nidditch. Oxford: Oxford University Press.
Losee, J. (1992). *A Historical Introduction to the Philosophy of Science.* Oxford: Oxford University Press.
Loux, M. L. (2002). *Metaphysics: A Contemporary Introduction.* London: Routledge.
Lowe, E. J. (1992). ‘Experience and its Objects,’ in T. Crane (ed.), *The Contents of Experience.* Cambridge: Cambridge University Press.
Luntley, M. (1999). *Contemporary Philosophy of Thought: Truth, World, Content.* Oxford: Blackwell.
Lycan, W. G. (2006). ‘The Meaning of Water: An Unsolved Problem,’ in E. Sosa & E. Vallanueva (eds.), *Philosophy of Language*, *Philosophical Issues*, 16, 184-199.
—. 2000. *Philosophy of Language: A Contemporary Introduction.* London: Routledge.
Lynski, L. (1977). *Names and Descriptions.* Chicago: University of Chicago Press.
Mackie J. L. (1980). *The Cement of the Universe: A Study of Causation* (1974) Oxford: Oxford University Press.
Maddy, P. (1990). *Realism in Mathematics*. Oxford: Oxford University Press.
Magee, B. (1997). *Confessions of a Philosopher*. New York: Random House.
Malcolm, N. (2001). *Wittgenstein: A Memoir.* Oxford: Clarendon Press. ‘Wittgenstein’s Philosophical Investigations.’ *The Philosophical Review*, 63, 530-559.
—. (1954). ‘Wittgenstein’s Philosophical Investigations.’ *The Philosophical Review*, 63, 530-559.
Malinowski, B. (1989). ‘The Problem of Meaning in Primitive Languages,’ published as a supplement in C. K. Ogden & I. A. Richards: *The Meaning of Meaning* (1923). Orlando: Mariner Books.
Manetti, G. (1990). *Theories of the Sign in Classical Antiquity.* Bloomington, IN: Indiana University Press.
Marcus, R. B. (1993). *Modalities: Philosophical Essays.* Oxford: Oxford University Press.

Maurin A-S. (2013). 'Trope Theory.' *Stanford Encyclopedia of Philosophy* (internet archive).
—. (2007). 'Infinite Regress – Virtue or Vice?' In *Hommage À Vlodeck: 60 Philosophical Papers Dedicated to Vlodeck Rabinowicz* (Lund University; internet resource).
—. (2002). *If Tropes.* Berlin: Springer Verlag.
McDowell, J. (1998). 'On the Sense and Reference of a Proper Name,' in *Meaning, Knowledge and Reality.* Cambridge, MA: Harvard University Press.
—. (1992). 'Putnam on Mind and Meaning,' *Philosophical Topics* 20, 1, 1992, 35-48.
McCulloch, G. (1989). *The Game of the Name.* Oxford: Oxford University Press.
McGinn, C. (2000). *Logical Properties: Identity, Existence, Predication, Necessity, Truth.* Oxford: Clarendon Press.
—. (1993). *Problems in Philosophy: The Limits of Inquiry.* Oxford: Blackwell.
—. (1976). 'On the Necessity of Origin.' *Journal of Philosophy*, 73, 127-135.
Mellor, D. H. (1977). 'Natural Kinds.' *British Journal for the Philosophy of Science,* 28, 299-312.
Mill, J. S. (2002). *System of Logic: Ratiocinative and Inductive* (1891). Honolulu: University Press of the Pacific.
—. (1979). *An Examination of Sir William Hamilton's Philosophy and of the Principal Philosophical Questions Discussed in his Writings* (1889), in J. M. Robson (ed.), *The Collected Works of John Stuart Mill* IX. London: Routledge and Kegan Paul.
Miller, A. (2007). *Philosophy of Language.* London: Routledge.
Misak, C. H. (2013). *The American Pragmatics.* Oxford: Oxford University Press.
—. (1995). *Verificationism: History and Prospects.* London and New York: Routledge.
Moore, G. E. (1959). 'A Defence of Common Sense,' in *Philosophical Papers.* London: George Allen & Unwin.
—. 'The Meaning of Real' (1953), in *Some Main Problems of Philosophy.* London: George Allen & Unwin.
Morris, C. W. (1938). *Foundations of the Theory of Signs.* Chicago: University of Chicago Press.
Mosteller, T. M. (2014). *Theories of Truth: An Introduction.* London: Bloomsbury.

Mulligan, K., P. Simons, B. Smith (2006). 'What is Wrong with Contemporary Philosophy?' *Topoi* 25, 63-67.
Mulligan, K., P. Simons, B. Smith (1984). 'Truth-Makers.' *Philosophy and Phenomenological Research*, 44, 297-321.
Neale, S. (1990). *Descriptions.* Cambridge, MA: MIT Press.
Neurath, O. (1931). 'Soziologie im Physicalismus,' in his *Gesammelte philosophische und methodologische Schriften.* Vienna: Holder-Pischler Tempsky.
Nietzsche, F. (1977). *The Portable Nietzsche*, ed. & transl. by W. Kaufmann. New York: Penguin Books.
Nishimoto, S. et al. (2011). 'Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies.' *Current Biology*, 21.
Nozick, R. (1983). 'Knowledge and Skepticism' In *Philosophical Explanations*. Cambridge, MA: Belknap Press.
Ostertag, G. (ed.) (1998). *Definite Descriptions: A Reader.* Cambridge, MA: MIT Press.
Papineau, D. (2012). *Philosophical Devices: Proofs, Probabilities, Possibilities, and Sets.* Oxford: Oxford University Press.
Patzig, G. (1980). 'Satz und Tatsache,' in *Tatsachen, Normen, Sätze: Aufsätze und Vorträge.* Stuttgart: Reclam.
—. (1977) 'Husserl on Truth and Evidence,' in J. N. Mohanty (ed.), *Readings on E. Husserl's Logical Investigations.* Den Haag: Nijhoff.
Pears, D. (1970). *Wittgenstein.* London: Penguin Books.
Peirce, C. S. (2005). *The Essential Peirce: Selected Philosophical Writings (1867-1896), vol. I.* Bloomington, IN: Indiana University Press.
—. (1931). *Collected Papers of Charles Sanders Peirce*, 8 vols. Vols. 1-6 eds. C. Hartshorne & P. Weiss. Cambridge, MA: Harvard University Press (1931-1935). Vols. 7-8 ed. A. Burks. Cambridge, MA: Harvard University Press (1958).
—. (1905). 'Issues of Pragmaticism,' *The Monist*, 15, 4, 481-499.
Perry, J. (1979). 'The Problem of the Essential Indexical.' *Nous*, 13, 3-20.
—. (1977). 'Frege on Demonstratives.' *Philosophical Review* 86, 474-97.
Pessin, A., S. Goldberg (eds.) (1996). *The Twin Earth Chronicles: Twenty Years of Reflection on Hilary Putnam's 'The Meaning of '"Meaning".'* New York: M. E. Sharpe.
Pitcher, G. (ed.) (1966). *Wittgenstein: The Philosophical Investigations: A Collection of Critical Essays*. Doubleday-Anchor.
Plantinga, A. (1974). *The Nature of Necessity.* Oxford: Oxford University Press.

Plato (1996). *The Collected Dialogues of Plato.* Harvard: Harvard University Press.
Polger, T. W. (2004). *Natural Minds.* Cambridge MS: MIT-Press.
Popper, K. (1972). *Objective Knowledge.* Oxford: Oxford University Press.
—. (1969). *Conjectures and Refutations.* London: Routledge.
—. (1959). *The Logic of Scientific Discovery*. London: Hutchinson. First published as *Logik der Forschung* (1935). Tübingen: Mohr Siebeck.
Price, H. H. (1953). *Thinking and Experience.* Oxford: Hutchinson University Press.
Putnam, H. (1989). *Representation and Reality.* Cambridge, MA: MIT Press.
—. (1981). *Reason, Truth and History.* Cambridge: Cambridge University Press.
—. (1975). 'The Meaning of "Meaning",' in *Mind, Language and Reality: Philosophical Papers, volume 2.* Cambridge: Cambridge University Press.
Quine, W. V-O. (1980). 'Two Dogmas of the Empiricism,' in *From a Logical Point of View*. Cambridge, MA: Harvard University Press. First Published in *Philosophical Review* 1951, 60 (1), 20-43.
—. *Word and Object* (1960). Cambridge, MA: MIT Press.
—. (1948/9). 'On What There Is.' *Review of Metaphysics*, 21-38.
Rasmussen, J. (2014). *Defending the Correspondence Theory of Truth.* Cambridge: Cambridge University Press.
Recanati, F. (2012). *Mental Files.* Oxford: Oxford University Press.
—. (2010). *Truth-Conditional Pragmatics.* Oxford: Oxford University Press.
—. (2004). *Literal Meaning.* Cambridge: Cambridge University Press.
Reichembach H. (1953). 'The Verifiability Theory of Meaning.' In *Readings in the Philosophy of Science*, ed. H. Feigl & M. Brodbeck. New York: Appleton-Century-Crofts.
—. (1938). *Experience and Prediction.* Chicago: Chicago University Press.
Reid, T. (1967). *An Inquiry into the Human Mind and the Principles of Common Sense*, in *Philosophical Works.* Hildesheim, Germany: Georg Olms Verlagsbuchhandlung, vol. 1.
Rescher, N. (1973). *The Coherence Theory of Truth.* Oxford: Clarendon Press.
Richardson, J. T. (1976). *The Grammar of Justification: An Interpretation of Wittgenstein*. London: Chatto and Windus, Sussex University Press.

Rosenthal, D. M. (2005). *Consciousness and Mind.* Oxford: Oxford University Press.
Rosh, E. (1999). 'Principles of Categorization,' in E. Margolis & S. Laurence (eds.), *Concepts: Core Readings.* Cambridge, MA: MIT Press.
Runggaldier, E. (1985). *Zeichen und Bezeichnetes: Sprachphilosophische Untersuchungen zum Problem der Referenz.* Berlin: Walter der Gruyter.
Russell, B. (2014). *The Autobiography of Bertrand Russell*. London: Routledge.
—. (1966). *An Inquiry into Meaning and Truth.* London: Allen & Unwin.
—. (1957). 'Mr. Strawson on Referring.' *Mind*, New Series, 66, 385-389.
—. (2009). *ABC of Relativity* (1925). London: Routledge.
—. (1929). *Introduction to Mathematical Philosophy* (1919). London: George Allen & Unwin.
—. (1994). 'The Philosophy of Logical Atomism' (1918), rep. in *Logic and Knowledge: Essays 1901-1950* (1956), (ed.) R. C. Marsch. London: Routledge.
—. (1910). 'Knowledge by Acquaintance and Knowledge by Description.' Proceedings of the Aristotelian Society 11, 108-128.
—. (1980). *The Problems of Philosophy* (1912). Oxford: Oxford University Press.
—. (1905). 'On Denoting.' *Mind, New Series*, 14, 479-493.
Ryle, G. (1957). 'The Theory of Meaning.' In *British Philosophy in the Mid-Century*, ed. C. A. Mace. London: George Allen & Unwin.
—. (1990). *The* Concept *of Mind* (1949). London: Penguin Books.
—. (1931/2). 'Systematically Misleading Expressions'. *Proceedings of the Aristotelian Society*.
Sainsbury, M. (1979). *Russell.* London: Routledge & Kegan Paul.
Salmon, M. (2002). *Introduction to Logic and Critical Thinking.* Belmont, CA: Wadsworth.
Salmon, N. U. (2005). *Reference and Essence* (1980). New York: Prometheus Books.
Salmon, W. C. (1966). *The Foundations of Scientific Inference.* Pittsburg: University of Pittsburg Press.
Saussure, F. (2005). *Cours de Linguistique Générale*, ed. C. Bally *et al.* (1916). Genève: Arbre D'Or.
Schlick, M. (1938). *Gesammelte Aufsätze 1926-1936.* Vienna: Gerold & Co.

—. (1938). 'Form and Content: an Introduction to Philosophical Thinking.' In *Gesammelte Aufsätze 1926-1936*, 151-249. Vienna: Gerold & Co.
—. (1936). 'Meaning and Verification.' *Philosophical Review*, vol. 45, 339-368.
—. (1910). 'Das Wesen der Wahrheit nach der modernen Logik.' *Vierteljahrschrift für wissenschaftliche Philosophie and Soziologie* 30, 386-477.
Schwarz, S. (2002). 'General Terms, and Rigidity: A Reply to LaPorte.' *Philosophical Studies* 109, 265-277.
—. (1978). 'Putnam on Artifacts.' *Philosophical Review* 87, 566-74.
Searle, J. R. (2015). *Seeing Things as they Are.* Oxford: Oxford University Press.
—. (2008). 'The Unity of the Proposition,' in Searle, *Philosophy in a New Century.* Cambridge: Cambridge University Press.
—. (2004). *Mind: A Brief Introduction.* Oxford: Oxford University Press.
—. (1999). *Mind, Language and Society: Doing Philosophy in the Real World.* London: Weinfeld & Nicholson.
—. (1998). 'Truth: A Reconsideration of Strawson's View,' in L. E. Hahn (ed.): *The Philosophy of P. F. Strawson.* Chicago: Open Court.
—. (1983). *Intentionality*. Cambridge: Cambridge University Press.
—. (1969). *Speech Acts: An Essay in the Philosophy of Language.* Cambridge: Cambridge University Press.
—. (1967). 'Proper Names and Descriptions,' in Paul Edwards (ed.), *The Encyclopedia of Philosophy.* London: Macmillan Publishers, vol. 6.
—. (1958). 'Proper Names.' *Mind* 67,166-173.
Sellars, W. (1962). 'Philosophy and the Scientific Image of Man,' in *Frontiers of Science and Philosophy*, Robert Colodny (ed.) (Pittsburgh, PA: University of Pittsburgh Press): 35-78.
Shields, C. (2007). *Aristotle.* London: Routledge.
Sidelle, A. (1989). *Necessity, Essence, and Individuation: A Defense of Conventionalism*. Ithaca, NY: Cornell University Press.
Simons, P. (1995), 'Meaning and Language,' in B. Smith & D. W. Smith, *The Cambridge Companion to Husserl.* Cambridge: Cambridge University Press.
—. (1994). 'Particulars in Particular Clothing: Three Trope Theories of Substance.' *Philosophy and Phenomenological Research*, 54, 553-75.
Skyrms, B. (2000). *Choice and Chance: An Introduction to Inductive Logic*. Wadsworth.
Sluga, H. (2011). *Wittgenstein.* Oxford: Willey-Blackwell.

Smith, A. D. (2005). 'Natural Kind Terms: A Neo-Lockean Theory.' *European Journal of Philosophy*, 13, 70-88.
Soames, S. (2003). *Philosophical Analysis in the Twentieth Century.* Princeton, NJ: Princeton University Press, vol. 2.
Sokolowski, R. (2000). *Introduction to Phenomenology.* Cambridge: Cambridge University Press.
—. (1974). *Russellian Meditations.* Evanston, IL: Northwestern University Press.
Stalnaker, R. (1978). 'Assertion,' in P. Cole (ed.), *Pragmatics*. New York: New York Academic Press, vol. 9, 315-32.
Stegmüller, W. (1989). *Hauptströmungen der Gegenwärtigen Philosophie, Band 1.* Stuttgart: Kröner Verlag.
Stenius, E. G. (1981). 'The Picture Theory and Wittgenstein's Later Attitude to it,' in I. Block (ed.), *Perspectives in the Philosophy of Wittgenstein.* Oxford: Oxford University Press.
—. (1960). *Wittgenstein's Tractatus: A Critical Exposition of its Main Lines of Thought.* Oxford: Blackwell.
Sterelny, K. (1983). 'On Natural Kind Terms.' *Pacific Philosophical Quarterly* 44, 110-25.
Strawson, P. F. (1991). 'Reply to Searle,' in L. E. Hahn (ed.): *The Philosophy of Strawson.* Chicago: Open Court.
—. (1987). *Scepticism and Naturalism: Some Varieties.* London: Routledge.
—. (1971). 'The asymmetry of subject and predicates,' in *Logic-Linguistic Papers.* London: Methuen.
—. (1991). *Individuals: An Essay on Descriptive Metaphysics* (1959). London: Routledge.
—. (1952). *Introduction to Logical Theory.* London: Methuen.
—. (1950). 'Truth.' *Proceedings of the Aristotelian Society,* Supplementary vol. XXIV, 1-23.
—. (1950). 'On Referring.' *Mind* 59, 320-344.
Stroll, A. (2009). *Informal Philosophy.* Lanham, MD: Rowman & Littlefield.
—. (1996). *Sketches and Landscapes.* Cambridge, MA: MIT Press.
Stroud, B. (1984). *The Significance of Philosophical Scepticism.* Oxford: Oxford University Press.
Swinburne R. G. (1975). 'Analyticity, Necessity and Apriority.' *Mind* 84, 225-243.
Tarsky, A. (1944). 'The Semantic Conception of Truth.' *Philosophy and Phenomenological Research*, 4, 341-375.

Teeple R. C., Caplan J. P., Stem T. A. (2009). 'Visual Hallucinations: Differential Diagnosis and Treatment.' *Primary Care Companion to the Journal of Clinical Psychiatry*, 11 (1), 2009, 26-32.

Textor, M. (2011). *Frege on Sense and Reference.* New York: Routledge.

Tononi, G. (2011). 'An Information Integration Theory of Consciousness,' *BMC Neuroscience*, 5-42

Tugendhat, E. (1992a). 'Überlegungen zur Methode der Philosophie aus analytischen Sicht,' in *Philosophische Aufsätze*. Frankfurt: Suhrkamp.

—. (1992b). 'Die Bedeutung des Ausdrucks "Bedeutung" bei Frege,' in *Philosophische Aufsätze.* Frankfurt: Suhrkamp.

—. (1976). *Vorlesungen zur Einführung in die sprachanalytische Philosophie.* Frankfurt: Suhrkamp. English translation: *Traditional and Analytical Philosophy: Lectures on the Philosophy of Language (Cambridge Philosophy Classics)* (2016). Cambridge: Cambridge University Press.

—. (1970). *Die Wahrheitsbegriff bei Husserl und Heidegger*. Berlin: Walter de Gruyer.

Tugendhat, E. & U. Wolf (1983). *Logische-Semantik Propädeutik.* Stuttgart: Reclam.

Urmson, J. G. (1983). *Berkeley.* Oxford: Oxford University Press.

Urmson, J. O. (1956). *Philosophical Analysis: Its Development between the Two Wars.* Oxford: Clarendon Press.

Velmans, M. & S. Schneider (2007). *The Blackwell Companion to Consciousness.* Malden, MA: Blackwell.

Vision, G. 2004. *Veritas: The Correspondence Theory and its Critics.* Cambridge, MA: MIT Press.

Waismann, F. (1986). *Logik, Sprache, Philosophie.* Stuttgart: Reclam.

Walker, R. C. S. (1988). *The Coherence Theory of Truth: Realism, Anti-Realism, Idealism.* London: Routledge.

Weber, M. (2010). *Die protestantische Ethik und der Geist des Kapitalismus* (1905). Munich: Beck Verlag.

Wiggins, D. (2001). *Sameness and Substance Renewed.* Cambridge: Cambridge University Press.

—. (1995). 'The Kant-Frege-Russell View of Existence: Toward the Rehabilitation of the Second Level View,' in W. Sinnot-Armstrong, D. Raffman, N. Asher (eds.), *Modality, Morality and Belief. Essays in Honor of Ruth Barcan Marcus.* Cambridge: Cambridge University Press.

Williams, D. C. (1953). 'The Elements of Being.' *Review of Metaphysics*, vol. 7 (2), 3-18, 171-192.

—. (1942). *The Ground of Induction.* Cambridge, MA: Harvard University Press.
Wisdom, J. (1969). *Philosophy and Psychoanalysis.* New York: Barnes & Noble.
Wittgenstein, L. (1984a). *Über Gewissheit, Werkausgabe Band 8.* Frankfurt: Suhrkamp.
—. (1984b). *Zettel*, *Werkausgabe Band 8.* Frankfurt: Suhrkamp.
—. (1984c). *Philosophische Untersuchungen, Werkausgabe Band 1.* Frankfurt: Suhrkamp.
—. (1975). *The Blue and the Brown Books.* Oxford: Oxford University Press.
—. (1974). *Letters to Russell, Keynes and Moore.* Ed. With an Introduction by G.H. Von Wright assisted by B.F. McGuiness. Oxford: Basil Blackwell.
—. (1984d). *Philosophische Grammatik, Werkausgabe Band 4.* Frankfurt: Suhrkamp.
—. (2001). *Wittgenstein's Lectures: Cambridge 1932-1935*, (ed.) Alice Ambrose. New York: Prometheus Books.
—. (1980). *Wittgenstein's Lectures: Cambridge 1930-1932,* (ed.) D. Lee. Oxford: Oxford University Press.
—. (1984e). *Philosophische Bemerkungen, Werkausgabe Band 2.* Frankfurt: Suhrkamp.
—. (1984f). *Ludwig Wittgenstein und der Wiener Kreis, Werkausgabe Band 3.* Frankfurt: Suhrkamp.
—. (1984g). *Tractatus Logico-Philosophicus, Werkausgabe Band 1.* Frankfurt: Suhrkamp.
Wrenn, C. (2015). *Truth.* Cambridge: Polity Press.
Zemach, E. (1986). 'Putnam's Theory on the Reference of the Substance Terms.' *Journal of Philosophy* 73, 116-27
Ziff, P. (1960). *Semantic Analysis*. Ithaca, NY: Cornell University Press.
Ziman, J. M. (1968). *Public Knowledge: An Essay Concerning the Social Dimension of Science.* New York: Cambridge University Press.
Zink, P. (1963). 'The Meaning of Proper Names.' *Mind* 72, 481-489.