

The topic of the dissertation is the reference of proper names. It criticizes the amalgam of the current standard theory, the direct-causal theory of reference and defends the Fregean sense theory of reference. In addition, the Fregean view is *neuronaturalized* by Gerald Edelman's idea of recurrent networks, combined with temporal synchrony of neural processing, which is generalized in the dissertation to explain conceptual information.

The direct-causal theory's main arguments - modal, epistemological and semantic - are all shown to be weak and based on mistaken presuppositions, like synonymies between proper names and descriptions. Causal communicative links are shown to be too weak for reference retention; it is also argued that intentions to retain the original referent are too weak. Moreover, it is suggested that the "intuition arguments" invoked by direct-causal theorists are misguided. They are shown to be based on a *basic schema of reference*, a postulation proposed in the dissertation to explain the precondition of referential use of language.

Reference is grounded on information transmission and its embodiment in neurocognitive processes, mainly having to do with neural memory systems. So contrary to the direct-causal theorists' view, reference is cognitive through and through, as Fregean theory maintains. Classical problems in philosophy of language, like the informationality of identity statements, propositional attitude contexts, *de re vs. de dicto*, among others, are solved.

ISBN 978-952-10-9860-4
ISBN 978-952-10-9861-1 (pdf)
Helsinki 2014
Unigrafia

Ari Peuhu
Towards a Neuroscientific Theory of Reference

Ari Peuhu

Towards a Neuroscientific Theory of Reference

**Philosophical Studies from the University
of Helsinki 44**

Filosofisia tutkimuksia Helsingin yliopistosta
Filosofiska studier från Helsingfors universitet
Philosophical Studies from the University of Helsinki

Publishers:
Theoretical Philosophy
Social and Moral Philosophy
Philosophy (Swedish)
P.O. Box 24 (Unioninkatu 40 A)
00014 University of Helsinki
FINLAND

Editors:
Panu Raatikainen
Tuija Takala
Bernt Österman

Ari Peuhu

Towards a Neuroscientific Theory of Reference

Academic dissertation to be publicly discussed, by due permission of the Faculty of Arts at the University of Helsinki, in auditorium XV, on the 7th of May, 2014 at 12 o'clock.

ISBN 978-952-10-9860-4
ISBN 978-952-10-9861-1 (pdf)

Contents

1. INTRODUCTION	1
General remarks.....	1
The Fregean sense theory of reference	4
The direct-causal theory	5
Overview	6
2. THE FREGEAN SENSE THEORY AND THE DESCRIPTION THEORIES OF REFERENCE	10
The motivation of Frege's theory of sense and reference.....	10
Fregean senses.....	11
Solutions to the puzzles	12
On Frege's explication of the senses.....	15
Russell's theory of descriptions	17
The cluster view of the description theory of reference.....	20
The relationships between Fregean sense and description theories of reference.....	24
3. THE DIRECT-CAUSAL THEORY OF REFERENCE	29
Background	29
The directness of the reference of proper names	30
Sense ₁ , sense ₂ and sense ₃	31
Rigid designation	34
The main arguments against the Fregean sense and the description theories of reference.....	36
<i>The modal, epistemological and semantical argument</i>	36
<i>Twin Earth and content variance argument</i>	39
<i>Kripke's semantical arguments from ignorance and errors</i>	40
The causality of the reference of proper names.....	42
<i>Kripke's picture of reference</i>	42
<i>Devitt's theory</i>	44
4. CRITIQUE OF THE DIRECT-CAUSAL THEORY OF REFERENCE	50
The modal and the epistemological argument	50
The semantical argument	58
<i>The status of the argument</i>	58

<i>Arguments from ignorance and error</i>	62
<i>Intentions to co-refer, reference borrowing by abilities to refer and the poverty of the causal chains</i>	63
<i>The meanings of proper names</i>	74
<i>Proper names and the basic schema of reference</i>	79
The Twin Earth argument.....	89
<i>Senses do determine referents</i>	89
<i>Epistemic symmetry principle</i>	92
<i>Internalism and senses₁₋₃</i>	97
<i>The (partly) non-conceptual status of the senses</i>	99
<i>The contextuality of the senses</i>	101
The content variance argument.....	106
Rigid designation and naming.....	111
Conclusions	116
Appendix: Devitt on the "four puzzles"	120
5. ACTUALIZED DESCRIPTIONS.....	135
Background	135
The actualized descriptions view	135
Soames' argumentation.....	139
Conclusion.....	142
6. FROM NEO-FREGEANISM TO NEURONATURALISTIC FREGEANISM	149
Problems with neo-Fregeanism	149
Recanati's account	152
The senses and grasping them.....	157
Wettstein's argument.....	160
Concluding remarks	163
Appendix: Wettstein's anti-Fregean revolutionary semantics.....	163
7. AN ATTEMPT AT NEUROCOGNITIVE THEORY OF REFERENCE	175
On methodology	175
Neurocognitive information.....	177
<i>"Cognitive" and "information"</i>	177
<i>From sensory to cognitive information through some functional neuroanatomy</i>	179
Recurrent/re-entrant networks	183
<i>Neuroscientific background</i>	183

<i>Edelman's TNGS</i>	185
Schemas	191
Individual concepts	194
Binding by temporal synchronization	198
The neurocognitive framework for reference.....	201
8. THE EXPLANATORY POTENTIAL OF THE NEUROCOGNITIVE THEORY.....	210
The classical problems	210
<i>The identity statements</i>	210
<i>Disambiguation</i>	212
<i>Propositional attitude contexts</i>	213
<i>Empty names</i>	217
<i>The existential statements</i>	221
<i>De re versus de dicto reference</i>	222
<i>Donnellan's distinction</i>	225
<i>Anaphora</i>	227
<i>Possible world semantics</i>	232
9. REFERENCE	239
On neuroscientific reference	239
BIBLIOGRAPHY.....	243

ABOUT THE WORK

This work was started sketchily towards the end of 1990s. Professor Gabriel Sandu thought as I did that some of my writings contained sufficient ideas and material for PhD. Then first attempt was arrogant “I-know-you-don’t” version at 2002. After that more humble attitude and detailed argumentation led to this version.

It has been long coming but here it is at last. Thanks are due to Gabriel Sandu and Panu Raatikainen for comments and advices, as well as to Paavo Pylkkänen. Paavo also made the publication of the work a reality when I was in dire straits. Thank you! Also Matti Sintonen arranged financial support earlier.

Most of all thank you Eeva for your love and being unforgiven companion.

I dedicate this work to Eeva, to my mother and to the memory of my father.

1. INTRODUCTION

"Having disregarded irrelevant objections they ought then to proceed to the most rewarding task of developing their point of view *in detail*, to examine its fruitfulness and thereby to get fresh insight, not only into some generalities, but into very concrete and detailed processes. To encourage such development from the abstract to the concrete, to contribute to the inventions of further idea, this is the proper task of a philosophy which aspires to be more than a hindrance to progress." (Feyerabend 1981: 161)

General remarks

Consider an utterance made by David: "Saul is hilarious". It is a simple utterance in surface, by which David refers to Saul and says something about him. But if the utterance and the multitudes of its likes really were simple ones, we would not be likely to prompt a simple question "Saul who?" To whom, or what, David refers, among all those who are named "Saul"? To answer this question is, in its barest, what the theory of the reference of proper names is all about. The proposed answers have been many and varied, which is a firm indication that there is nothing simple about the reference of proper names. But the current received view seems to be, at least within the analytic philosophy of language, that the direct-causal theory (as I call the theoretical package, see chapter 3) is superior to the Fregean sense theory of reference (the description theory of reference understood to be part of the latter). According to the former theory a proper name refers because there are causal links from the naming event of the referent to the later communicative uses of its name. The latter theory explains reference of proper names with senses, cognitive representations that characterize the referents uniquely. But having said this it must be immediately noted that the matter is not so clear-cut. The neo-Fregeans, beginning with Michael Dummett's studies, have put forward their accounts of the reference of proper names (as well as demonstratives and indexicals) which differ to varying extents from the direct-causal theory, both with respect to the presuppositions and details. And that is not all, for some neo-Fregeans accept some of the central tenets of the direct-causal theory (for instance Gareth Evans and Francois Recanati). As to myself I will be finding a niche towards the Fregean end of this continuum, but I also acknowledge the few sound points of the direct-causal theory, albeit mostly in attenuated forms. But on the whole, to lay open the dialectical situation, currently the proponents of the direct-causal theory of reference form the majority of the reference theorists. We can say that the direct-causal theory of reference has become the paradigm of reference research in analytic philosophy.

But as we will see that consensus is hardly anything else than an illusion, albeit firmly believed one. Accordingly I have two main aims in this work. The first is divided into two further subaims. The first of these is to argue that the arguments the direct-causal theorists have given against the sense theory are not sustainable; the second is to argue, meshed with the former arguments that the direct-causal theory will not do as an explanatory theory of the reference of proper names. This is the promised

disillusionment. This argumentation will also show that the common way of speaking about “the Fregean description theory of reference” that the direct-causal theorists haplessly cultivate is improper and has only contributed to its own widespread acceptance. (Though there is one recent exception: Howard Wettstein writes “In *Naming and Necessity*, Kripke criticizes the ‘Frege-Russell description theory,’ which perhaps neither Frege nor Russell held precisely in the form described...” (Wettstein 2004: 97); maybe the tide is turning from expository convenience to accuracy in substance?) To my view this is mildly put: one corollary of my argumentation is that neither Frege nor Russell held Kripke’s version of the “description theory” at all. Plainly, Obviously, Frege held *sense theory* that is *not* equivalent to any version of the description theory. Russell’s theory is different in its tenets from the version of the description theory Kripke found fault with, which is *Searle’s* (or Searlean) cluster theory. As to the second main aim of this study, it is to develop the Fregean perspective into a neuroscientific framework, or neurocognitive theory of the reference of proper names (and by obvious implication, a framework for reference in general). At this point it must be mentioned that the neo-Fregeans have never equated the Fregean senses with definite descriptions. As Harold Noonan remarks, mentioning Dummett’s and Evans’ works (Dummett 1973, 1981), (Evans 1982), “...we must not suppose that [Frege] thought that sense *must* be descriptive, and in so far as modern critics assume this they are attacking a straw man.” (Noonan 2001: 180) So am I merely wasting my efforts and time in rearguing an old hat? The answer is no because the assimilation of Frege’s (and the Fregean) view to the description theoretic view is still widely held by the direct-causal theorists. Moreover, and as Noonan has also pointed out, it is still much debated issue whether or not Kripke’s and the other direct-causalists’ arguments really hit on the assumed Fregean target (Noonan 2001: 214). And because the direct-causal theory is the prevailing paradigm in the current reference research the argumentative misconceptions that have led to the widespread assimilation of that view need to be disentangled. (It also seems to me that many of the misconceptions have not been uncovered before.) The second and related point is that the actual situation, the tug-of-war between the direct-causal theory and the (neo-)Fregean theory is much more dramatic and detrimental to the former theory: as I will argue at length the direct-causal theory is at most, and at best, only a subtheory of the Fregean theory; Frege already had all the makings the direct-causalists have claimed to be their innovations.

Let me also mention right here some of the more specific conclusion I will arrive at. It turns out that the causal connections from naming of the referent to the later uses of that name do not make up the main explanatory factor behind the reference of proper names; the neurocognitive information that the causal links carry is what explains reference. Likewise the intentions to co-refer, the intentions to keep on referring by the hearer to what the speaker is referring – reference borrowing - by his use of a proper name as such, are explanatorily forceless. The reason is again that the cognitive and informational factors are left crucially out of the picture. And the intuitions used in the direct-causal theorists’ arguments, i.e. that there still is a genuine reference when all the information about the bearer of the name in question is incorrect, are explained away by showing that they are misconceived: they are not primarily referential-semantic at all but indicate a neurolinguistic-cum-perceptual-cum-functional structure underlying all referential uses of language (“the basic schema of reference”). As for the positive contributions of this work I argue that the information that explains the reference of proper names is tailor made for the Fregean sense theory of reference -

as the neo-Fregeans have realized. In developing my version of the general Fregean approach I need only to correct some faults in the most promising neo-Fregean attempts along my way towards the neuroscientific framework of the reference of proper names. It is in order to point out already here that the usual charge launched by the direct-causal theorists to the effect that the Fregeans have missed the communicative factor in reference retention is also an unfortunate misconception. The Fregean sense theorists just have not emphasized the communicative factors in so many words, thus not the causal factors on which the direct-causal theorists have focused. That is mainly because the neo-Fregeans have thought that the communicative factors are obvious to the referential use of language. (I think that that should have been evident from the examples they have used. But obviously it has not.) In a nutshell my general argument against the direct-causal theory of reference is that it has seemed to be the correct theory - hence has become widely accepted - because what explanatory power it has it in fact owes to the Fregean sense theoretic factors inherent in it but not recognized as such so far. As I said at most the direct-causal theory is a subtheory of the Fregean sense theory of reference. There is an irony here, which is that although the direct-causal theorists have stressed the general point, they have largely remained silent on the very particular factors that retain the references of proper names in the communicative exchanges, i.e. the informational-cum-cognitive factors.

Once we get rid of the simplified picture and assumptions behind the direct-causal theory and bring forth the neural and cognitive dynamics of the brain, we are well on our way to developing a unified theory of reference. So in addition to presenting a rather long critical examination of the direct-causal theory this work attempts to give us such a general but sufficiently detailed theory of the reference of proper names, my other main aim. The theory I propose is an interdisciplinary one. The reason is that I do not see how otherwise we could have any hope of explaining the reference of proper names (and reference in general). After all the referential use of language is what we are immersed in frequently; it is a form of human communicative action. And just as other forms of human actions it depends on the many different structures and functions of our brains, in relation to our surroundings that are informationally represented in our brains. Consequently it seems to me that it is hopeless to believe that we could have an adequate explanatory theory of reference without taking into account the relevant pieces of knowledge ranging from neurobiology to cognitive sciences. This obviously implies that no "purely philosophical" theory of reference has any chance. Reference is an empirical phenomenon, consequently it should be studied also with the methods of science and not only with the methodological preconceptions prevalent in philosophy (like "intuitions" and folk semantically charged examples). Among the latter I also count in the Kripkean assumption that a philosophical theory proper is only such an account that gives necessary and sufficient criteria for the phenomenon it studies; or put in other words, that the Fregean sense theory of reference is committed to giving synonymy analyses of proper names by the definite descriptions providing necessary and sufficient conditions for the senses, conceived of as meanings, of proper names. (As is well known Kripke himself doubted the prospects of purely philosophical theory of reference, i.e. necessary and sufficient conditions theory. A hypothesis as to why he demanded that methodology from the Fregean and description theories, thus violating the argumentative parity, will be offered in section 4.1.) Contrary to all that the interdisciplinary perspective is the big picture from which my attempts spring forth.

The ism that is usually given to this kind of orientation is *naturalism*. Briefly, naturalistic philosophy uses scientific results and theories quite extensively in its attempts to provide answers to philosophically conceived problems. This involves giving up some old-fashioned projects, for example epistemology that tries to justify knowledge claims by basing them on some undeniably secure foundations. Or, as just alluded to and more relevant to our subject, demands such as strong semantic equivalences like synonymy between proper names and the senses as (linguistic) meanings (expressed by the definite descriptions). Still, the philosophical and the scientific endeavours proceed in parallel and in interactive manner in our quests for ever better descriptions and explanations of the natural phenomena. In a way the kernel of this naturalistic approach can be expressed concisely by the slogan that good philosophy is not philosophy. That is, not only philosophy, and in some cases not even mainly.

The Fregean sense theory of reference

According to the Fregean sense theory of reference referential expressions have senses (*Sinne*) related to them such that those senses determine uniquely their referents. It is true that Frege was rather niggardly in his explication of the senses of proper names – even though proper names - *Eigennamen* - is a wide category for him: it includes what we take to be the ordinary proper names, but also the definite descriptions as well as declarative sentences, or assertions for Frege) Ontologically speaking Frege's view was clear, however. Senses are abstract "third realm" entities that all competent speakers of a language can "grasp". They are not subjective entities like psychological ones, though the graspings are psychological acts. This assures us of the objectivity of the senses, hence of the objectivity of communication and science, i.e. the efforts to find out truths. But we will see later that the common assumption by the direct-causal theorists that Frege equates the senses of proper names with the definite descriptions is seriously mistaken. I say "seriously" because that assumption has led to a bunch of all too easily refutational arguments against the Fregean sense theory. It seems that the misconception can be mainly explained by the fact that Bertrand Russell proposed a theory of the definite descriptions that holds that the ordinary proper names are disguised (telescoped, abbreviated) descriptions. Consequently it has been easy to assimilate Frege's theory to Russell's because Russell's views were well known to the later generations of the analytic philosophers of language. By the same token this explains the primary place given to the "cluster view" proposed by John Searle. But the assimilation is mistaken. It should be kept firmly in mind that for Frege the senses are the ways the referents are presented and singled out (to be thought and talked about). As Frege says, metaphorically, senses "illuminate" the referents, and, not so metaphorically, they are the modes of presentation of the referents. Hence there is no exclusive commitment to the definite descriptions as totally equivalent to the senses. Still it is as obvious that the definite descriptions *express* in most cases at least partially the senses of proper names. That was evident also to Frege: among the few examples he gave of senses there are descriptorial expressions (as in the famous "Aristotle note" (Frege 1997a: 153, n. B)).

As I argue Frege's senses, and my neurocognitive Fregean senses, are not exclusively descriptive, even though it is obvious that parts, or elements, of them can be linguistically expressed by the descriptions. (I will be saying more on this in places

throughout this work.) Just because of this and because – as Frege held – we could not grasp the senses except by the help of language, the danger of mutual assimilation of the sense and the description theories has loomed large. (To that testifies the rather common way of talking about “Fregean description theory” and, indeed, assimilating the versions of the latter to the former.) The reason I dwell on this is that I realize that along the way I have become to take it as obvious that Frege’s approach and the *Fregean* approach (as it is partly exemplified also in the neo-Fregean studies) are not description theoretic, with the consequence that I have not taken sufficiently literally the common view of “Fregean description theory”. (But nevertheless I also defend the description theory indirectly as well as directly (chapter 5).) That is, my arguments against the direct-causal theorists’ counterarguments to Frege’s and the Fregean theory, understood as perhaps inadvertent sorts of the description theories, attempt to show that they fail one way or another. And because the direct-causal arguments were intended to sink the description theory also – to repeat: the distinction between Frege’s, the Fregean and the description theory (and *its* versions) were not clear and/or much cared about (and still do not seem to be) – it obviously follows that my defence of Frege’s and the Fregean views subsumes the defence of the description theory in the appropriate places.

The direct-causal theory

A rival to the sense theory of the reference of proper names came to market in the late 1960s and in the beginning of the 1970s. The proponents of that theory argue against the Fregean theory, i.e. what they think is Frege’s theory, that descriptions are neither necessary nor sufficient in determining the referents of proper names they are related to. (These arguments have been extended to the natural kind terms, but I have to restrict my sweep here.) In particular the direct-causal theorists have argued that proper names are not synonymous with the definite descriptions associated to them, i.e. that the descriptions do not give the *meanings* of proper names. At most the descriptions can be used to *fix* the referents, but later on they might become semantically irrelevant to the referential function of the very names. (As a foretaste of my criticism this *assumes* that the Fregean senses are meanings, but that is only partly correct. (See section 4.2.4.)) Moreover it assumes a very stringent conception of the purpose of semantics: meanings as giving synonymy relations. That assumption becomes *non sequitur* the way the direct-causal theorists have used it, as we will also see. According to the direct-causal theory proper names are mere “tags” and “nonconnotative appellations” and they refer without the semantic mediation of any Fregean senses. Another way to express this general tenet of the theory is that proper names are *rigid designators*: contrary to the definite descriptions they refer to the same actual individual in all counterfactual, or modal, contexts (possible worlds). Because the definite descriptions do not function exactly as proper names do, they are neither referentially, nor in general semantically, equivalent to proper names. So the direct-causal theorists argue that the directness of reference of proper names amounts to there being no Fregean referent-determining senses. Instead according to the causal component of the theory the referent is given a name and all (or most) of the subsequent uses of that name, with the intention to co-refer, refer to that “target” because of the causal-cum-communicative links mediating the connection from the target to the very uses of that name. The causal-cum-communicative component of that theory purports to provide the explanatory mechanism, and the minimum of

physicalistic respectability to the view (for there is no *direct* reference; that would be a metaphysical oddity).¹

It should be mentioned at the outset that because this is a work on the reference of proper names I am not going to discuss demonstrative and indexical expressions. But I see no obstacles to developing a neuroscientific theory of reference of indexicals and demonstratives in the spirit of the Fregean approach. One result that these studies should support is that the often-advertised arguments to the effect that the Fregeans could not account for the demonstratives and indexicals are not sound. The arguments are based on the assumption that it almost defines a Fregean that according to him all referring expressions must be given descriptonal equivalents. But in fact the very anti-Fregean arguments (from the "essential indexical" *à la* John Perry) are admirably Fregean in showing that there are more to the senses than what meets the descriptonal eye.

Overview

In chapter 2 I will present Frege's theory of sense and reference; not all of it but those elements that are relevant to this study. No new observations are made with respect to that theory; only few remarks in places where I have thought that something needs to be said of correcting little misunderstandings, or to shed a light on some direction that has been neglected so far. In chapter 3 the same is done to the direct-causal theory of reference, except that I say even less anything new. The main arguments against the Fregean sense theory of reference are also presented there: the modal, epistemological and semantical argument, and the arguments from Twin Earth and content variance of the senses.

The critical stance against the direct-causal theory is launched in chapter 4. every one of the first three main arguments is shown to be a failure. The modal and the epistemological arguments are refuted quite quickly. They have to do with the assumption behind the direct component of the direct-causal theory, according to which proper names are not semantically equivalent to the definite descriptions expressing the Fregean senses. But the semantic assumption, synonymy, and the epistemological assumption, apriority, which are behind the arguments, are not part of the Fregean sense theory and never were. The semantical argument, which has mainly to do with the causal element, has grown quite intricate during the years, so the counterargumentation takes the main portion of the chapter. I will argue that the use of the causal-cum-communicative links, or chains, could not explain the reference of proper names; only information about the referents can do that. I also argue that the intentions to co-refer are forceless in referent retention; and that the referential "intuitions", commonly invoked when the descriptions about the purported referent that are shown to be false in the counterexamples the direct-causal theorists have amassed, are misinterpreted by the proponents of the theory. The plain consequence of this is that when a sense of a proper name does not apply to a unique individual that proper name does not refer uniquely, which is just the verdict of the sense theory. During the course of the argumentation I will also defend the Fregean theory, especially with respect to the particular nature of the Fregean senses. We will see that they are not nearly as "conceptually pure" as the anti-Fregeans have made them look like. On the contrary, we will see that the senses are so rich that the direct-causal

theory, in places where it does have some explanatory adequacy, has already been presupposed by the Fregean theory.

The Twin Earth argument fails because it involves too strong an epistemic assumption: it violates what I call "the epistemic symmetry principle". In short, and in general, the Twin Earth scenarios violate that principle because they take different epistemic perspectives to be semantically equivalent, our current perspective and that of the Twin Earthians. But that needs to be argued for, not assumed. The content variance argument states that the senses the speakers associate with a single proper name vary to the extent that they could not be part of the semantic values of proper names, therefore the modes of presentation of the referents are not semantically relevant. But this argument presupposes a relativization of the uses of proper names, and that presupposition amounts to the individuation of the referents of proper names in question. And that individuation can be attained only with the Fregean means, not by causal ones. From this it also follows that the notion of rigid designation is sound but argumentatively impotent one because it in fact presupposes the individuation of the referents of the rigid designators in question (contrary to Kripke's claim that transworld identity, i.e. individuation, puts the individuating cart before the referential horse). So all in all, if all the arguments are correct, the direct-causal theory is at best only a subtheory of the Fregean theory, for the former theory fails in all other respects than those having to do with the *preconditions* of reference, which obviously the Fregean theory contains. The direct-causal theory is surprisingly empty of substantial explanatory and referential content of its own.

In chapter 5 the actualized descriptions view is defended. I will argue that it is the most natural reading of the description theory and that Frege, Russell and Strawson (but not Searle, it seems) held it at least tacitly. Moreover I argue that if this reconstruction is correct, Kripke's modal argument involves a crucial slip of failing to distinguish between two readings of the definite descriptions, the actualized one from the predicative. The result is that, contrary to the almost unanimously held belief, the description theory of reference interpreted correctly has also been immune from the direct-causal attack all along.

In chapter 6 I take up some problematic features of the current neo-Fregean views, especially their reliance on what I call "secretarial analogies", i.e. "files" and "dossiers" as the reconstructions of the Fregean senses. I correct some of the most promising ones and in this way lead the neo-Fregean approach to the domains of the cognitive and the neurosciences, which are put in full use in chapter 7. In chapter 6 I also explain in the Fregean terms the distinction between proper names as rigid designators and the definite descriptions as non-rigid, i.e. when they are not used as actualized-rigidified devices of reference. (The ground for this is already laid in section 4.5.) In constructing a framework for reference of proper names I use two seminal explanatory notions in the current theoretical neuroscience, i) re-entrant/recurrent network architecture, after Gerald Edelman's theory of neuronal group selection, and ii) the binding of information by temporal synchrony in those networks. I also reconstruct in neurocognitive terms the old notion of *individual concept* by (reductively) identifying it with the notion of the schema about an individual. These I explain by a further reduction to the neural level entity in Edelman's theory, a *global mapping*. Naturalistic as my perspective to philosophy and the theory of reference (and semantics) is, the framework is an exercise in

neuronaturalizing the reference of proper names. In chapter 8 I put the framework into use and offer solutions to the classical and most pressing problems in the philosophy of language and mind. In the last chapter I discuss the general features and aspects of the neuronaturalistic theory of reference, and also counter some skeptical arguments from the indeterminacy of reference.

Why I defend not just the Fregean theory or just the description theory but both – albeit the latter, so to say, as a by-product of the defence of the Fregean theory? Have not both of the theories been decisively refuted? Well, if I would believe that, I would not defend them, would I. The only diagnostic answer that the question deserves, to my mind, apart from the obligatory detailed argumentation to be given in this essay, is that the question is only based on a misguided consensus. For the situation is seems to be the following. The original arguments by Donnellan and Kripke, and by Putnam to some extent, were good, but as I will show, they were such only because they were based on such alleged assumptions and conceptions about the theories they attacked that the arguments could not have shown their case. And because the misconceptions have passed unnoticed and in effect have become accepted by the majority, the consensus has been formed. Contributing factor has been also that the occasional papers noticing the serious weaknesses of the direct-causal theory – like the early ones (Loar 1976) and (Schiffer 1978) – have been mostly ignored, so the consensus has grown (in good faith in most cases, no doubt), the result obviously being that the Fregean and the descriptive theory have not really received the re-hearing they deserve. So my primary aim is to defend and develop a modern version of the Fregean view.

In concluding this introduction I want to take up shortly an issue I have quite often been confronted with. The neuroscientific framework I develop has, for some people, seemed quite out of place in a work that is supposed to belong to philosophy. But that seems to me to be an erroneous opinion. The motivation for the neuroscientific (and cognitive scientific) approach is very philosophical indeed: as I see it no adequate explanatory account could be achieved without taking relevant sciences into account. This is *metaphilosophical* motivation for it seems to me that the reference research has been too much in the grip of the old-fashioned philosophical methodology (for instance counterexamples and strong semantic presuppositions in conceptual analyses). To offer an alternative to this in action, a *neuronaturalistic* account of reference as a representative of *philosophical naturalism*, is what this essay is premised on.

Moreover I do not see how my approach differs from that kind of analytic philosophy in which tools from mathematics and logics as special sciences are used (sometimes very heavily). Science is science, formal or substantial. That kind of approach is wholly accepted as a methodology to try to solve philosophical problems, so what argument could there be to the effect that the neuroscientific approach is not just as acceptable? And, mind you, the most common way of practising even analytical philosophy by giving examples as evidence from everyday situations – especially in reference research – belongs to that kind of practice. The data is not from science but it is offered as data anyway, so what could be supposed to be the difference here? I do not see any. Keeping philosophy “pure” is what suits old-fashioned philosophers but not those of us seeking explanations of foundational problems, philosophical included.² I see my neuronaturalism as one specimen in the recent science-informed

philosophical studies of virtually all kinds: for example philosophies of matter, of space and time, of the status of species, of thinking and perception and, indeed, of consciousness have all received remarkable boosts by opening up to what scientists are saying about those subjects. It is time we heed what the relevant sciences are saying about reference.

Notes

1. Some writers speak about "causal-historical reference" or "historical explanation theory" of reference (for example Donnellan). But this reflects the difference of focus on the particular level or components of the direct-causal theory, not a difference in explanatory substance.
2. I have always thought that there is something profoundly unphilosophical in the demand for only "purely" philosophical answers – and ways to handle – philosophical problems. (See note 2 in chapter 6.) To my experience much of that kind of philosophy – conceptual analysis – tends mostly to be a waste of time – though some of it presents admirable as specimens of intellectual exercises. But even then one gets the impression of clever tricks without passion and vision of a larger background; nothing of worth could be built on that, I surmise. The most extreme, or to my mind preposterous, claim I have encountered is that it is all very well to explain with the help of science how information is processed in the brain, but that does not answer the "philosophical" problem about what makes the neural-cum-cognitive structures and processes *represent* the external objects. But on what grounds is it presupposed that the scientific theories do not explain *precisely that problem*? When the scientific story of the processes from external stimuli through neural-cognitive representations to behavior have been told, what room is there left for any genuine "philosophical" answers? I see none. In other words I could not see anything distinctively philosophical about the problem of representation and reference to external objects. Traditionally the problem has been philosophical but that was so only because the scientific means were not yet in any very advanced state. (If one means, or conflates this to, the problem of skepticism about the external world, he is in different ball park, for then the scientific story could not be accepted to begin with, for it more or less presupposes the existence of the external world.) But conceptual analysis is not the only villain here. In addition to its influence there has prevailed also the epistemic-ontological one, the old hankering for "transcendental" perspective. It has quite much to do with the "pure philosophy". According to this line of thought for every empirical domain of study there are non-empirical preconditions of its soundness and validity. These have to be non-empirical for otherwise they would be only contingent and incapable of having the desired *a priori* status. The misconception behind this view is that it does not see that empirical disciplines, working as interactive parts of the whole enterprise of knowledge, could be self-reflective in the way that their methods and concepts can apply to themselves and be checked by both more general and specific disciplines (and these checked in their turn as they develop). In contrast to this I hold that philosophical reflective activity can be caught up with scientific and empirical means.

2.THE FREGEAN SENSE THEORY AND THE DESCRIPTION THEORIES OF REFERENCE

The motivation of Frege's theory of sense and reference

In addition to his interest in formal logic Friedrich Ludwig Gottlob Frege was also deeply interested in thoughts (*Gedanke*) and senses (*Sinne*). His interest in language was only of derivative nature, but, as he pointed out more than once, it is mainly through language that we are able to investigate thoughts and senses. The obvious reason is that linguistic expressions are the primary vehicles of expressing thoughts and senses. This is all evident in Frege's writings (but it is also convincingly argued by Tyler Burge (Burge 1979)).¹

Consequently it is only natural that one of the targets, and at the same time test cases, of Frege's theory were some puzzling linguistic statements, especially the puzzles created by identity statements.² The most famous of them is that which manifest differences in the cognitive content or cognitive significance (*Erkenntniswert*) of the respective expressions. In most cases it is trivial to state that "a is a" (in the formal mode), but it may be informative to state that "a is b" even when the referent of "a" is the same as the referent of "b" (that is, when they have the same *Bedeutung*). Thus it is usually trivial to say that Hesperus is Hesperus, meaning that the planet called "Hesperus" is self-identical (or something as trifle as that). But to state that Hesperus is Phosphorus is informationally valuable in that the statement asserts that the two expressions refer to the same object (the planet Venus, as it happens). The ancient astronomers were most likely somewhat surprised to discover that the star (the planet, as it happens) in a particular place and route across the evening sky is in fact the same star as that with the particular location and route in the morning sky. The identity statement "Hesperus is Phosphorus" was not a piece of linguistic triviality or tautology for them.

The notion of cognitive significance needs to be explained, especially when we are concerned with the reference of proper names. The Millian theory, of which the direct-causal theory is *the* representative, and which is the current main competitor of the sense theory, takes the sole meaning, the sole semantic value, of a referring expression to be the referent of that expression. In the case of proper names the meaning of the name, the semantic value, is that individual to whom it refers, be that a person, an animal, a city, or whatnot. On the Millian theory the identity statements have the same meaning, the same semantic value, if their referents are the same. (This is a curious way to express the point, for there is in that case only one referent; but the point is clear.) In the above example from the domain of the ancient astronomy the meaning of the identity statement "Hesperus is Phosphorus" is that the planet Venus is Venus. But, as is widely acknowledged, this view has the consequence that it is unable to explain – without quite drastic changes of what is considered to be only semantically relevant – why, for instance, a person might not believe that Hesperus is bright although that same person believes that Phosphorus is bright. The reason for this heavy theoretical hedging is that according to the prevailing Millian view – or neo-Russellian with singular propositions – the person *does* also believe that Hesperus is bright (Salmon 1986). But this just flies on the face of the obvious cognitive fact that when asked the person himself would be likely to deny this belief. Thereby the epicyclic moves by the Millians.

Fregean senses

Frege's aim was to develop a theory that both solves the puzzles and reveals the primary role of the thoughts and the senses over the natural language expressions expressing them more or less accurately. To this end Frege introduced the distinction between *Sinn* and *Bedeutung* in his famous paper *Über Sinn und Bedeutung* (Frege 1997a). The theory is based on quite a rich semantical framework, but those aspects of it that relate to the puzzles are so well known that we can present them in a compact way. Frege distinguishes *Sinn* and *Bedeutung* for three categories of linguistic expressions. In general *Sinn* is the sense of an expression (its content or meaning in the current jargon); *Bedeutung* is its referent (denotation, designation). Sense is a functional as well as a cognitive-epistemic notion. To further specify the nature of the sense Frege says that it involves the way or the mode of presentation of the referent (*die Art des Gegebenseins*). Thus in particular the sense of an assertoric sentence (or a statement, an utterance) is the thought (*Gedanke*) it expresses. Thought in this sense is what is meant by the current notion of proposition a sentence or an utterance expresses. For example the sentence "the velocity of light is constant in vacuum" expresses the thought that light has constant velocity in vacuum. It is a different thought from the one the sentence "light travels 299 792 458 meters per second in vacuum" expresses, which states the precise value of the velocity of light in vacuum. The thoughts are similar in content but not the same because their respective senses differ, but *Bedeutung* is the same, the constant value of the velocity of light in vacuum. The important point to realize is that for Frege thoughts as senses of assertoric sentences are the guarantors of the objectivity of communication. This requires of the thoughts that they form the common stock for all competent users of (a particular) language. Frege did not hold the view that thinking always requires language. He thought it possible that there could be creatures that can grasp thoughts and senses without the mediation of any language, but we humans are incapable of grasping and communicating thoughts without the mediation of language. Thoughts as well as senses are third realm entities in Frege's framework. That is, they could not be perceived, neither are they subjective ideas (*Vorstellungen*) in the consciousness of a thinker. Because of this abstract nature of the thoughts and senses, and their being available for all competent users of language, they are fit for objective scientific theorizing and semantics; ideas, images and such occurants are not because of their subjectivity. Neither are the latter entities in any way involved semantically in senses and thoughts: in contrast to a subjective idea an objective thought "needs no owner" (Frege 1997b: 337).

The sense of a proper name is, then, the way the referent is given or presented to a thinker. For example the expression of a sense of the name "Aristotle" may be "the pupil of Plato and the teacher of Alexander the Great" (Frege 1997a: 153, n.B). This sense gives the referent of the name "Aristotle" (in contrast to the sense, say, "the famous Greek shipping magnate", which refers to Aristotle Onassis). It should be strongly emphasized that in this same place Frege admits that different persons may have different senses of the same proper name. In regard of the sense of "Aristotle" just mentioned, Frege continues that "Anybody who does this will attach another sense to the sentence 'Aristotle was born in Stagira' than will someone who takes as the sense of the name: the teacher of Alexander the Great who was born in Stagira." The possible multitude of different senses of one and the same proper name (syntactically

or formally individuated not to beg any semantic questions) is a characteristic feature of these expressions in natural languages. But that the variance of the senses ought not to exist in the *ideal* scientific language was one strong motivation of Frege's studies. It is easy to be inclined to think that, as the senses are intersubjectively shared "meaning entities" in Frege's theory, people share all the senses of a particular expression. But that need not be so. It is better to allow that senses are *shareable*, not that they are all actually shared. For example if we shared all the senses of a particular proper name, there would not arise any doubts as to what or who is the referent of that name. But such doubts do arise occasionally, so the view that the senses are shared rather than sharable would seem to lead to the conclusion that we are infallible in our knowledge with respect to individuals (because we would also have access to all true thoughts about individuals). But this does not follow from Frege's view either, and he in fact denied any such consequence: "The sense of a proper name is grasped by everybody who is sufficiently familiar with the language or totality of designations to which it belongs: but this serves to illuminate only a single aspect of the *Bedeutung*, supposing it to have one. Comprehensive knowledge of the *Bedeutung* would require us to say immediately whether any given sense attaches to it. *To such knowledge we never attain.*" (Frege 1997a: 153; emphasis mine).

The third category of expressions with the senses is the concept (or predicate) words. They are somewhat more complex semantically in that they have senses but their *Bedeutungen* are the concepts themselves. That is because Frege countenanced concepts (functions in general) in addition to objects in his ontology. Concepts have extensions in their turn. These are the object or objects falling under the particular concepts. In this essay I am not concerned with the Fregean concepts, neither with the distinction between concepts and objects, so I will not continue on this matter further (but see (Frege 1997c,d) for details).

Solutions to the puzzles

Let us see how Frege's theory solves the well-known puzzles mentioned above. The puzzle of informational identity statements is solved in the following way. The sense of a proper name is that which gives or singles out the referent of the name. This is to be understood as a specification of the way the particular sense a speaker grasps represents the referent, so that one person can refer to the planet Venus by invoking the sense, i.e. the thought, expressed by the sentence "the star at the morning sky", and another person by the sentence "the star at the evening sky". As this example shows one and the same entity can be referred to by many different senses. And it can also be referred to with a not wholly correct sense: the star in question is a planet, Venus. This explains why the identity statement "the Morning Star is the Evening Star" is informative and not trivial. The Morning Star/Evening Star example shows also that the Fregean approach is *not* a version of the description theory, for – as just pointed out – the respective descriptions expressing the respective senses are incorrect, but the senses are still about the same object. So there is at least one more element implicit in the senses because they pick up the *correct* referent.

One can also be oblivious as to the sameness of the referent because one has at one's disposal only senses, and the two senses expressed by "the Morning Star" ("Phosphorus") and "the Evening Star" ("Hesperus") are different. In general the senses explain the cognitive fine-grainedness of the thoughts. In contrast the coarse-

grainedness that characterizes the Millian approach, just because of that coarseness, is not able to explain the different cognitive significances of the trivial and informative identity statements. (It has been argued, for example by Nathan Salmon (1986) that the Millian approach, or in effect the Russellian theory, does not wreck on the rock of the informative identity statements. But all these kinds of solutions need to postulate something that accounts for the different cognitive roles of the thoughts expressed by the identity statements, therefore in effect they only work because of the Fregean nature of those postulations, like the guises under which something is believed in Salmon's theory; see section 4.3.2, note 20.)

Frege's theory works also well in solving the puzzle the propositional attitude contexts are prone to manifest. (Here I use believing as the representative of the large class of the propositional attitudes.) One may believe that Hesperus is not bright even if he believes that Phosphorus is bright. The reason is that the senses (the thoughts) the two sentences express differ from each other. As there is nothing in the sense of "Hesperus" that would indicate that it has the same referent as "Phosphorus", it becomes natural that one may believe something about Hesperus without believing that same thing about Phosphorus - and *vice versa* for the sense of "Phosphorus".

Frege elaborated this solution a little further. Expressions in the propositional attitude contexts refer to their *customary senses*, not to their customary referents (*Bedeutungen*) that are the objects they refer to in direct speech, like the planet Venus in our example. The customary senses become the indirect *Bedeutungen* in the propositional attitude contexts. In other words, in contrast to the direct speech contexts, we are talking primarily about the thoughts of the persons, not about the referents that are specified by the senses that are parts of those thoughts.

In the same manner we can also explain why two central principles of extensional logic fail in the propositional attitude contexts. Substitutivity of identicals states that co-referring terms can be substituted *salva veritate* for one another, i.e. without a change in the truth-value of what the sentences say of which the substitutable terms are parts. In the propositional attitude contexts the referents are not the same because the referents are the senses (or thoughts), and they may differ from one another. The other principle, existential generalization, states that from "a is F" can be inferred "there is an individual such that it is F" (formally: "Fa" => " $\exists x Fx$ "). However one may believe that Santa Claus is an old man (=Fa) without Santa Claus existing. This apparent violation of the existential generalization is explained in the same vein: one's belief is about the sense expressed by the sentence "Santa Claus is an old man", but it is not about any existing object, about any customary *Bedeutung*.³

Frege's theory is not restricted to what we currently understand by proper names. Among them (*Eigennamen*) Frege included also the definite descriptions. But, as we saw, the theory is designed to work as well with sentences: a thought has a sense, or more precisely, the thought *is* the sense of the sentence expressing it. The truth-value of the thought (True or False, *tertium non datur*), thus derivatively of the sentence that expresses it, is its *Bedeutung* (as the object a proper name refers to is the *Bedeutung* of the name). Frege's treatment of sentences differentiates those sentences that are involved in finding out truths about reality from those sentences that are of an epic or a fictional nature – or "mock thoughts" as they are usually called. The latter do not have truth-values as their *Bedeutungen* because some (or all) of their constituents lack

Bedeutungen. Thus "Odysseus was set ashore at Ithaca while sound asleep" is such a sentence, and accordingly it does not have truth-value because "Odysseus" has no referent in reality. The sentence expresses a sense, though. (For how else could we say that it belongs to the domain of epic?) But in spite of these type of sentences Frege held that proper names in our current sense, as well as definite descriptions and sentences are proper names with senses, even if they do not always have *Bedeutungen* (Frege 1997a: 157; 1997e).

It should be mentioned already here that it has become common to take the Fregean sense as that factor which *determines* the referent. There has been some controversy over the alleged incoherence in Frege's terminology related to this issue. If sense is the mode of presentation of the referent, then there presumably is no sense when there is no referent if the slogan "sense determines referent" is taken at its face value and *modus tollens* is applied. But to argue this way is to rely on the unlikely premiss that Frege would not have noticed the incoherence. I take it simply that this assumption amounts to a pragmatic *reductio* of the argument itself. Moreover, Frege himself is explicit on the point that proper names lacking *Bedeutungen* still have senses (Frege 1997a: 156-8). It seems to me that David Bell, Michael Beaney and Gideon Makin have already resolved the issue (Bell 1990:274-5; Beaney 1997: 22-4; Makin 2000: 112-17). They all argue against Gareth Evans' interpretation of Frege's approach with respect to empty names, i.e. senses that do not have referents. Evans charges Frege on inconsistency because Evans takes it that to speak about sense as the mode of presentation will not do if there is no referent. But as both Bell and Makin easily show such charge is groundless. Frege himself explicitly states in *Über Sinn und Bedeutung* (and elsewhere) that sentences have senses even when they have parts that have no referents. Frege also says that one is not assured of the referent of the sense. To my view this seems to be equivalent to holding that senses are genuine even when they do not have referents. And as Bell says "for Frege the sense of an expression is the *condition* that must be met by anything that is the reference of that expression...the possession by a sign of a determinate, coherent, and intelligible sense is entirely compatible with that sign's lacking a reference." (Bell 1990: 275; emphasis mine) According to Beaney senses as the modes of determination of the referent are the routes by which the referent is arrived at, the conditions that the referent must meet to be that very referent. This does not demand that there will always *be* a referent. It is true that the talk about the modes of presentation of the referents tends to be talk about those senses that *do* have referents. But that should not lead us to the hasty Evansian conclusion, for it is just that Frege tended to employ this latter explication of the senses because of his interest in truth and in the foundational logical matters where the senses have a central role and their *Bedeutungen* are guaranteed. Makin makes essentially the same point with the implication that we should not take the "route" metaphor too seriously here. At any rate "...no grave consequences for Frege's position ensuing from admitting that conceiving of sense as 'mode of presentation', though heuristically valuable, and true whenever there is a referent, cannot serve as a general characterization." (Makin 2000: 117) In short, then, it seems that Evans' argument is only a claim based on a too literal reading of "determining *the* referent". It is another question why Evans committed such a mistake. And here we have, I think, a deeper undercurrent to be uncovered. For it seems to be not very likely that Evans would have made such a mistake without a certain background conception. This conception might be what also determined his insistence on the "Russellian thoughts"/"de re senses", i.e. that when there is no object of the (purported) thought,

there is no thought. We see this same pattern here: the existence of the object – to be determined by a sense and to be thought about – is a precondition for Evans' views. But here I cannot try to argue for this hypothesis.

On Frege's explication of the senses

Frege did not say too much about what senses are (except, of course, that the sense of a thought is composed of the senses of its constituent senses). Frege used some metaphors occasionally to enlighten the nature of senses, for example that sense illuminates, *beleuchtet*, the referent from one side. We saw above one of the few occasions when Frege characterizes the senses of proper names: the senses of the name "Aristotle". Ontologically senses are abstract entities, hence not sensible objects and not objects of consciousness, what Frege called *Vorstellungen*. But of course there has to be some connection from sensibilia to the senses, for how otherwise could they have any role in thinking and language use (as well as in perception (Frege 1997b)). Frege left us little concrete clues as to how the connection is effected. He talked about "grasping" and "apprehending" senses and thoughts (verb: *ergreift*, *erfassen*; noun: *Auffassung*). He said that grasping or apprehending is a psychological act: "How does a thought act? By being grasped and taken to be true. This is a process in the inner world of a thinker which may have further consequences in his inner world, and which may also encroach on the sphere of the will and make itself noticeable in the outer world as well." (Frege 1997b: 344) This still leaves us very much in the dark because psychological acts are subjective but senses and thoughts are objective in Frege's view. (Although as such this does not yet help much, but it is understandable because Frege was not interested in psychological matters; he rather abhorred them, for he thought that psychologizing the issues was a sure way to perdition in any semantic and logical matters.)

Anyway we should dispel right here the rather common belief that Frege was a description theorist. He did hold that we humans could grasp and express the senses only with the help of language, but that there could be creatures that grasp the senses (and thoughts) directly. So it follows that the senses should not be equated with the definite descriptions. Moreover Frege held that both proper names (in our sense) and the definite descriptions are *Eigennamen* with their *own* senses. So equating the senses of the ordinary proper names with the senses of the definite descriptions (let alone the senses of the former with the descriptions themselves) is out of place. The direct-causal theorists may have failed to appreciate this point about the subsidiary role of language and consequentially been led to presume the equivalence of the senses and the descriptions because of their predetermined agenda in semantics. I will argue in section 4.1 that this indeed is so. Undoubtedly they were misled by the few examples of the senses Frege gave; but how else could he have given them other than by the linguistic means, by descriptions, when he held that we grasp and describe the senses only by linguistic means? However, it could be argued that the senses are weakly descriptive because parts of them, or elements of them, could be expressed by descriptions. And that is true, for – again – how else Frege could have given his few examples of the senses like the sense of "Aristotle" being "was born in Stagira and taught Alexander the Great". But this is trivial in light of what was just said: we humans could not access the senses but with the help of language. Moreover it can be argued that the senses are only *partially* expressible: they have elements that do not "yield" to specific linguistic characterization (in contrast to quantifying over those

elements and mentioning them by general terms). This will be argued in section 4.3.5 (Ateb-Afla example) and in the chapters 6 and 7 (and with the example of the twin blondes in section 8.1.1). Note by the way that I am not claiming that *every* sense of a proper name is such that it could not be completely linguistically articulated by descriptive means. No doubt there are, or at least could be, such senses, like descriptive names envisioned by Gareth Evans (“Julius” = “the one who invented the zip”).

It is evident from Frege’s writings that the senses have cognitive function and explain the cognitive significance of expressions. It should be noted that despite the current looseness in meaning between “psychological” and “cognitive” one should not be led to think that the psychological grasping of the senses and the cognitive contribution of them are epistemically or semantically related in Frege’s view. The difference in the cognitive values between, for example, the trivial and informative identity statements is a semantic fact that is explained by the difference of the senses the different referential expressions occurring in them – even though that only manifests itself in the linguistic and behavioral level. By the same token we can also explain what one knows: the total knowledge of a person is formed by the totality of the true thoughts he grasps. In other words Frege’s semantic theory of sense and reference is closely related to epistemic matters. But the psychological acts involved in “apprehending” and “grasping” thoughts and other senses are strictly distinct from these semantic, epistemic and cognitive aspects.

What was just said about thoughts and senses is rather compact presentation of that part of Frege’s theoretical apparatus. However it is more or less all there is to be said about the nature of the senses. However I will bring up couple of striking features of them later when I answer arguments directed against the general nature of the senses by the direct-causal theorists (sections 4.3.4 and 4.3.5). The barrenness of Frege’s elucidation of senses is disturbing, to be sure, but it is also a rewarding tenet. That is because it leaves one with an opportunity to try and develop both a more detailed and wider theory of senses in relation to a theory of reference. From this perspective the rather negative result of Frege’s elucidations become a motivating boost to modernize, to *naturalize*, the notion of Fregean sense. It also explains the seemingly different attitude that for instance Michael Dummett expresses when he says that “Frege tells us much about what senses are” and then goes on to give all the same formulations from Frege as were given above (Dummett 1981: 80). Presumably whether one is an optimist or a (qualified) pessimist on the issue of the sufficient characterization of the nature of the senses depends on one’s theoretical aims. It is true that Frege could have been even more silent on the issue. He might have left it just to mentioning that the senses of, say, “Morning Star” and “Evening Star” differ from each other in spite of the fact that they have the same referent. But from my perspective Frege’s characterizations of the senses are not sufficiently specific. I think that the notion of the sense is of utmost importance in semantics and in the theory of reference (as well as in the philosophy of mind). Because I am in this study after a modernization and neurocognitive naturalization of the senses with respect to the theory of reference, I welcome the opportunity to seize the senses in service of a scientific account of them.

Russell's theory of descriptions

There are two strands that motivated Bertrand Russell's celebrated theory of the definite descriptions (Russell 1905). It is a theory, or we should perhaps say, as is common, a representative piece of path-breaking philosophical methodology, which has since been emulated hundreds of times in analytic philosophy. The story is therefore a familiar one, but let me review it shortly. First motivation for Russell was that he became deeply dissatisfied with proposals like the one put forth by Alexius Meinong, which postulated (or hypostatized) all kinds of entities which "subsist" in order to explain the meaningfulness of statements, even incoherent and contradictory ones like "there is a square that is round". The ontological extremity of this approach can be readily seen from that very statement, for it presupposes some thing to deny the *existence* of that same thing (which most likely lies behind Meinong's motivation to postulate the mode of *subsistence* for entities). Generally speaking this Meinongian line is ready to postulate any kinds of entities and modes of being (*sic?*) to explain anything, not just semantic and linguistic puzzles. The ontological carelessness that Meinong's approach brings with it became not to suit Russell's ontological taste. He wanted instead to get clear about the workings of natural language by the use of formal logic, not to take natural language at its face value in order to make ontological leaps. In essence Russell wanted to show that if a linguistic expression can be given a paraphrase, or some such semantical formulation, the expression could be dispensed with without any ontological crimes being committed.

The second motivation was primarily epistemological. Russell maintained that there is a distinction between knowledge by acquaintance and knowledge by description. Roughly, the former knowledge is about entities that are the direct objects of one's experience. Although there has been debate about whether sense data or common material objects are the primary direct objects, I think that for Russell both are included (if not, nothing depends on that here). Thus one's sensations, emotions, perceived objects and such are known by acquaintance. Also included were at least some memories and awareness of oneself, and presumably universals as the referents of common nouns, predicates and relation words. In natural language there are expressions that refer to these items of knowledge by acquaintance, expressions what Russell called "logically proper names". They refer directly to the entities they are about; thus their meanings are the very entities they refer to. The class of logically proper names consists of expressions like "this X", "that X", "now", "here", "I", and such; expressions for one's sensory contents and perceived objects ("this dog").⁴

Russell's claims that the sentence "X exists" is a tautology and sentence "X does not exist" is a contradiction might seem a little hard to understand. But if we keep in mind Russell's epistemic view and its accompanying thesis that only immediate sensory "objects" can be acquainted with, the claims become fathomable. When "X" is a logically proper name of such an object it is about a particular element of one's experience. Therefore it is indeed tautologous to say that X exists, for "X" is the name of that non-recurring element of one's experience; "X" as a logically proper name occurs only once and is tied once and for all to that experiential element. Or, if it recurs, its name would still be "X": "X" is permanently undetachable from that experience. (This includes the memory of X because as a memory item it still could not be renamed, even if that seems so. The reason is that one would be naming the experience of remembering X.) In the same vein to say that X does not exist is a

contradiction because at the moment "X" is said it refers to an unique experiential element. (There could not have been a name for that experience any time before it occurred.) Maybe an analogy helps also. The utterance "I do not exist" is a contradiction because it could not be uttered without its utterer existing. Likewise "I exist" is tautology (in most cases at least) because there is no need to utter it when the utterer - obviously - exists.

If we accept the idea of logically proper names as expressions that could not fail to refer directly, the idea of the theory of the definite descriptions in its Russellian version is the following. What are called *ordinary* proper names, or just proper names in the current jargon, do not belong to the class of logically proper names. The ordinary proper names and the definite descriptions are incomplete symbols: they are meaningful only in a context. In contrast logically proper names are capable of standing alone because they cannot fail to refer: there is always an item to which such an expression refers. This idea opens up a way to analyze the statements that deny the existence of something, the negative existentials ("X does not exist"), without any need to postulate suspect entities allegedly to be the referents of the expressions. Russell's theory of the definite descriptions maintains that the ordinary proper names are "disguised" ("abbreviated", "telescoped") definite descriptions. This way to analyze them shows that they have quantificational structure. The reason for this treatment is that if the subject X in the statement "X does not exist" were a logically proper name, the statement would be contradictory because "X" as a logically proper name would have its referent as its very meaning, but by assumption there is not such referent (that is what the negative existential says). The statement is, however, perfectly understandable and does not commit its utterer to any contradictions or incoherences, therefore "X" is not a logically proper name.

To present the crux of Russell's theory in a compact form, let us analyze the statement "the round square does not exist" the way Russell encourages us. "Round square" is not a logically proper name, for presumably no one is acquainted with any round squares. This leads us to analyze the statement in a way that enables us to preserve its meaningfulness. The *analysans* that Russell would offer is "it is not the case that there is one and only one entity which is both round and square". Because there is no such entity that is both round and square (at the same time), the original statement is meaningful and true.

Expressions like "round square" are not the most likely expressions in the everyday discourses, so let us also see how Russell's theory analyzes a meaningful expression that has no referent (at the moment of its utterance). The famous example of the king of France will do here. How to analyze the statement "the present king of France is bald"? The problem is of course that there is no current king of France, therefore the expression is not a logically proper name. Because we are not allowed the easy way out of postulating a "subsistent" king of France, we have to find another way to explain the meaningfulness of the statement. This is offered by Russell's analysis that is based on the contention that the grammatical surface form is not the logical form of the statement. The latter form reads, after the analysis, "there is one and only one king of France, and he is bald". And this is false just because there is no present king of France at this moment. This way the analysis also preserves the law of excluded middle, i.e. that any declarative sentence is either true or false, *tertium non datur*. As Russell says, if you go through the bald entities you will not find the king of France

there, neither among the non-bald entities. This creates a threat to the law of excluded middle, but the analysis evades it and moreover does that by showing why the statement is meaningful: it has the form of a quantificational sentence. (The same analysis also handles the negative existential "the king of France does not exist" the same way as it dealt with the round square.)

The king of France example analyses a definite description, but what about the ordinary proper names? If the expression "Scott" were a logically proper name, thus always unfailingly referring, it would not be even meaningful to express doubts about Scott's existence by saying that "Scott does not exist". Let us look at the sentence "Scott is the author of *Waverley*", where "is" is the is of identity and not the is of predication. This sentence is not a tautology because it clearly expresses something that one may not have known before. The reason for this is, according to Russell, that "the fact that the one is a name and the other a description. Or they might both be descriptions." (Russell 1956: 247) If that were not the case, that is, if the description "the author of *Waverley*" were semantically equatable with a name, akin to a logically proper name, the sentence would be expressing a tautology because both "Scott" and "the author of *Waverley*" would refer to the same individual. To reveal this argument more completely let us suppose that "the author of *Waverley*" is a logically proper name. This means that there must be an individual, say X, such that the expression refers to it. From this it follows that the whole expression is then equivalent to "Scott is X". Now this leaves us with the alternatives: either "Scott is X" is false, in which case Scott is not identical to X and "Scott" and "X" do not refer to the same individual; or "Scott is X" is true and Scott is identical to X, in which case the statement is a tautology. Consequently the whole statement "Scott is the author of *Waverley*" is false or tautologous. But it is neither, so the assumption must be rejected. That is, "the author of *Waverley*" is not a logically proper name.⁵

It is rather seldom noted (or maybe even noticed) that the standard examples of the Russellian analysis of definite descriptions are incomplete. For example the analysis of "the present king of France is bald" is stopped at the stage of "there is one and only one king of France, and he is bald". But the whole intent of the theory of the definite descriptions is to analyse *all* denoting expressions to only quantifiers and predicate expressions (and identity sign plus logical connectives). Therefore "the king of France" should receive further analysis. (This point is clearly presented by Salmon (Salmon 1989).) But this full intention of the theory has proved difficult to fulfill. Or rather that it could not be fulfilled has been one of the charges the direct-causal theorists use against the description theory. I will come to that issue later. Here it must be mentioned that Russell supposed that the analysis does go through in all stages down to the logically proper names. i.e. with respect to all the ordinary proper names and descriptions. This supposition stems from one of his metaphysical views that he seemed to have held throughout his whole philosophical career. He argues that ordinary proper names embody a false metaphysics, the substance metaphysics (Russell 1948: 95-9). Russell argues instead for metaphysics in which the function of proper names is to name complexes of qualities and spatiotemporal relations that make up an individual. This characterization of Russell's view makes it evident that he did not give up, or that at least he kept coming back to, the doctrine of knowledge by acquaintance and the logically proper names being about the direct items of experience. The relationship of predicate expressions, many of which Russell took to express universals, to qualities is of course problematical, but to the extent there exists

such a relationship it becomes natural to suppose that the analysis of descriptions can be followed through. The motivation is that the completion of Russell's theory and its harmonization with his epistemology both require that. That is because according to Russell's view all linguistic expressions, *a fortiori* definite descriptions, must be reducible to constituents with which we are acquainted in experience.

It is not altogether clear that Russell's theory can solve all the puzzles he considered as its objectives (see note 5). But that is quite beside the point for my purposes because the main thrust of Russell's theory has survived to our days in one form or another. The crux of that theory has always been that proper names are abbreviated (or disguised) definite descriptions, when the descriptions individuate one and only one referent, i.e. are unique in their reference-determining function. If we drop both the atomistic metaphysics and epistemology of knowledge by acquaintance, a move unanimously adopted by the subsequent theorists, and as a consequence focus on the ordinary proper names, the main idea of the theory of the definite descriptions remains viable. It has been extended, so we better turn next to the most famous of the post-Russellian versions.

The cluster view of the description theory of reference

After Russell's original theory of definite descriptions the most important extension of the description theory was the so-called "cluster theory". Especially John Searle has argued for this view (Searle 1958). Peter Strawson is sometimes mentioned as a proponent of that view because of his account in his book *Individuals* (Strawson 1959). I do not concur with this – as we will see in places in chapter 4, so I will here discuss only Searle's position. The cluster theory is an extension of Russell's theory in that its central claim is that a proper name is typically associated with many different descriptions, not just one or several which are uniform across all the speakers. The theory states also that it may happen that not all of those descriptions contribute to the reference of proper names they are associated to, but only those descriptions contribute that best fit the referent, or those that are most important to the characterization of the referent. (Though Russell's theory can also be taken to imply quite naturally that the descriptions form conjunctions, not disjunctions.)

In the mentioned paper Searle takes an issue with the argument – later much used by the direct-causal theorists - that once a proper name is learned with the help of descriptions, the subsequent uses of it will be independent of those descriptions, and that they will not form part of the sense of the name in question because the name does not have a sense but the name only refers. This Searle finds implausible because, were we to discover that the factual knowledge about, say, Aristotle would turn out to be false, we would rather say "for this reason that Aristotle did not exist after all, and the name, though it has a conventional senses, refers to no one at all..." (Searle 1958: 168). Searle then continues by asking what is the difference between proper names and other singular referring expressions, definite descriptions in particular? His answer is that in contrast to the definite descriptions proper names do not specify any characteristics of the objects they refer to. But there has to be some connection between a proper name and its object. This Searle finds in the feature that the referring uses of proper names *presuppose* that their objects have certain characteristics, and they are expressible by descriptions. "Therefore, referring uses of 'Aristotle' presuppose the existence of an object of whom a sufficient but so far unspecified

number of these statements are true." (Searle 1958: 170-1) It is almost always left open which of these descriptive criteria, or characteristics, are those the speaker is relying on in a particular use of a proper name. But that is so because that is the very point of the difference between proper names and the definite descriptions: if the characteristics could be decided once and for all for all uses of a particular proper name, the resulting definite descriptions would take the role of the proper name. But because this is not so, the looseness of the connection between proper names and the characteristics of their bearers expressed, or expressible, by the definite descriptions is not only an acceptable but a desirable feature as well. As Searle himself puts it: "...the uniqueness and immense pragmatic convenience of proper names in our language lie precisely in the fact that they enable us to refer publicly to objects without being forced to raise issues and come to agreement on what descriptive characteristics exactly constitute the identity of the object. They function not as descriptions, but as pegs on which to hang descriptions." (Searle 1958: 172)

The looseness of the relation between proper names and the definite descriptions allows, then, not any strict characterization (except, maybe, in some very simple and secure cases like the above mentioned descriptive names). That explains it why one has to have recourse to such expressions as "best fit" to characterize that relation. But this should not give rise to any serious *argumentative* clashes because the expression "best fit" that I used above is in fact taken from an antagonist to the cluster view (and to the Fregean sense theory): it is used by Nathan Salmon in his presentation of the cluster view (Salmon 1982: 10).

In this connection it is in order to mention for comprehensiveness' sake that Wittgenstein had a very similar view in relation to the looseness of the descriptive criteria associated with proper names. Though he did not propose anything like a sufficiently developed account of the reference of proper names (that would presumably have been distasteful to him, judged from the general tone of his later philosophy), it has become a sometimes habit to quote what he says about the negative existentials. So I habituate myself also:

"If one says 'Moses did not exist', this may mean various things. It may mean: the Israelites did not have a *single* leader when they withdrew from Egypt-or: their leader was not called Moses-or: there cannot have been anyone who accomplished all that Bible relates of Moses-...But when I make a statement about Moses-am I always ready to substitute some *one* of these descriptions for 'Moses'? I shall perhaps say: by 'Moses' I understand the man who did what the Bible relates of Moses, or at any rate, a good deal of it. But how much? Have I decided how much must be proved false for me to give up my proposition as false? Has the name 'Moses' got fixed and unequivocal use for me in all possible cases?" (Wittgenstein 1953: 36-7)

Leonard Linsky draws from this quotation what seems to me to be its adequate referential moral (Linsky 1977: 86-7, 104-5). According to it proper names do not have a fixed senses for all cases of their uses (let alone for all possible cases). Still, they function well in our everyday communication in spite of the looseness, the lack of fixed overall senses. If I may add, Frege's famous Aristotle note is, in effect, another

way to highlight this when the looseness stems from the intersubjective variance of senses. So it seems to me that the looseness has nothing to do with the nature of senses but, so to say, with the *variability* of their ranges for different thinkers and speakers. Frege's view involves a natural generalization of the variability of the senses, for these ranges vary also *intrasubjectively* from an occasion to an occasion with respect to the same name. Empirically this can be understood that a person may forget some senses (or in modern terms: information) about the referent; or he may remember different senses at different occasions. (I am not indicating, let alone saying, that Frege would not have accepted this as a matter of course. He just did not mention the point – maybe because psychological acts were something of secondary interest to him at best.) There, then, Searle's and Wittgenstein's views meet each other and with Frege's. When someone says that "Moses does not exist" that simply means that he says that Moses does not exist, but there is no preconceived, or prefixed, set of descriptions related to "Moses" such that they would not change in any way when someone else says that "Moses does not exist" (in a different occasion, and maybe with different concerns).

When that is put in epistemic terms, as Linsky does (Linsky 1977: 105), it means that normally there are no facts, fixed in advance, which would lead us to state that NN does not exist. And sometimes it is left open whether NN exists or not, or did exist. (I think that such cases of undecidedness can be found in history.) The relevance of the emphasis on the epistemic matters is that it brings out in the open the parting of the ways of the sense (as well as the description theory) and the direct-causal theory of reference. The reason is that one main tenet of the latter theory that the reference of proper names does not depend on anything epistemic, or cognitive. Consequently even the looseness of the relation between the descriptions and proper names to which the former are related is an anathema to the direct-causal theorists. But to jump ahead a little, we will see that the direct-causal theory fails on the score of the empirical facts of the referential uses of language, even on its own terms (intentions to co-refer, abilities to refer explaining reference borrowings), and the crux of the failing is the disregard of the very epistemic and cognitive factors in the referential use of proper names. The lack of epistemic interest together with the empirical inadequacy leaves the theory with only (allegedly) intuitively supported arguments. But I will argue that even these are based on misunderstandings.

But let us go back to Searle's version. He claims that between a name and the description associated with it there is a logical connection: it is "necessary" that, for example, "Aristotle has the logical sum, inclusive disjunction, of properties commonly attributed to him: any individual not having at least some of these properties could not be Aristotle" (Searle 1958: 172). So it is not usually sufficient that one description singles out the referent, but a collection of them is sufficient to do that. Aristotle was the one and only one person who taught Alexander the Great, was Plato's most famous pupil, was born in Stagira, wrote *Nichomachean Ethics*, *Poetica*, and so on. However it seems to me that the necessity Searle invokes here is of rather innocent sort. Especially it is not a representative of any essentialism with respect to the properties of individuals. That not any essentialist consequences are intended is quite evident from Searle's characterization of the difference between proper names and the definite descriptions. Searle just leaves it open what descriptions could be involved in any particular use of a proper name. But even if Searle's notion of necessity is so strong that it implies some kind of essentialism then let it be so. Any untoward consequences

that this might have do not affect the Fregean theory that is the focus of my interest here.

Searle does not discuss the Fregean puzzles in his paper, but only gives the solution the cluster theory delivers to the negative existential statements "X does not exist". According to Searle the statement "Aristotle never existed" says, not that the proper name "Aristotle" has never referred to anyone, but that the statement asserts that a sufficient number of conventional presuppositions and descriptive statements associated to the proper name "Aristotle" are in fact false. However, we do not know precisely which of these are false because the precise conditions that constitute the criteria for the correct uses of "Aristotle" are not given by language (Searle 1958: 173).

It seems to me that the positive existential statements can be explained in the same way. Remember that the alleged problem with them was that according to Russell's account if a proper name "X" is a name, in the sense of being a logically proper name, it becomes tautologous, or trivial, to say that "X exists". According to the Searlean explication, as I see it, the statement "Aristotle exists" asserts that a sufficient number of the presuppositions and descriptions that are associated to "Aristotle" are in fact true, but which ones is not yet decided once and for all and presumably could not be, either. Therefore the existential statement is not a tautology.

To continue with my reconstruction of the deliverances of the cluster theory, the informativity of some identity statements is explained in a straightforward way also: some descriptions from the cluster are not either uniformly shared between different speakers, or one particular speaker does not relate them as characterizing the same referent. Thus it may come as a surprise to someone that Saloth Sar was Pol Pot because the descriptions he relates to "Saloth Sar", if he relates any, are different from those he relates to "Pol Pot", and the descriptions themselves do not say or imply that Saloth Sar was Pol Pot. And *vice versa* with "Pol Pot". Instead his presuppositions are most likely different, i.e. he takes the different sets of descriptions to characterize different referents. This same explanation takes care of the propositional attitude contexts, for a person may believe that Pol Pot was a cynical dictator but not believe that Saloth Sar was a cynical dictator, despite the fact that the names refer to the same person.

Before continuing any further it must be pointed out, to remove possible misunderstanding, that the textbook account of the cluster theory is too restricted in part (Devitt & Sterelny 1987: 43). Michael Devitt and Kim Sterelny allow that some descriptions may have greater weight than others, but they do not consider the fact that the weights of the distinct descriptions do change in different contexts. Thus when the topic is logic, "the systematizer of syllogistic logic" presumably weighs more than "the son of the court physician Amyntas II", with respect to "Aristotle". But when the topic is, say, history it may be that the second description weighs more. As far as I can see, this sort of contextual looseness is one factor contributing to the looseness between proper names and the definite descriptions that Searle emphasized. In light of this it is not an acceptable move to argue against the cluster theory by claiming that it is "committed to selecting some descriptions that *define* the name and rejecting others that express *merely accidental* properties of its bearer." (Devitt & Sterelny 1987: 44) Moreover, as Searle said and as certainly seems to be the case, we do *not* give definitions of proper names in ordinary practice of speaking (Searle 1958: 166).

To conclude this section I want to point out that the cluster version can be made even more loose than that only best fit or such notions are allowed. For it seems that even false descriptions can be used to refer to a certain individual, i.e. to the individual they are false about. For example the description "the first man who sailed around the globe" commonly refers to Fernao de Magalhaes. That this is false does not matter for the reference: Fernao de Magalhaes died at the island of Mactan in Cebu 27.4.1521, and the rest of his crew sailed with the last ship and arrived eventually at San Lucas in Spain 6.4.1522. Because of the many other facts known about the adventure and because the chief captains used to be in such esteemed positions in those days the reference of the description "the first who sailed around the globe" is taken to be sufficiently uniquely referential to F de M. (Otherwise we would state that the first to sail around the globe was Juan Sebastian de Cano, the captain of the remaining ship, but we, usually, do not.) So it seems that the cluster can allow even some false descriptions used *referentially* within it as long as there are other descriptions known and true of the referent. (That Columbus found America is perhaps the most common instance of the reference by false description.) I think that in general this proposal might deserve a further study, but I eschew from that study here.

The relationships between Fregean sense and description theories of reference

There are similarities and dissimilarities between Frege's and Russell's theories, as well as with the main extension of the latter, the cluster theory. In this section I will discuss some of them, albeit quite shortly. My motivation for the discussion is that it will turn out to be of utmost importance later in this work to have a clear picture of the crucial relationships between the sense and the description theories. This is especially so because it has become almost invariably accepted that the two theories are essentially the same in their content and so of their explanatory status. But this claim is incorrect, as I will argue.

We can begin with an underlying similarity between the Fregean sense and the description theories. Both Frege and Russell were motivated by a background conception, or interest, in constructing their respective theories. Frege's theory is largely based on his interest in thoughts coupled to his lifelong urge to construct the formal logic by which all true thoughts, in effect the whole of science, would be expressed. Tyler Burge has argued, and convincingly to my mind, that Frege's primary interest in developing his semantic theory was epistemological; Michael Dummett has also argued for this view in his works on Frege. Contrary to what is usually assumed, then, Frege was not a philosopher of language, except so to say derivatively. Natural language was of secondary importance to him but investigating thoughts and their role in finding out truth was his primary goal. This explains why Burge can focus on Frege's epistemological perspective: only the true thoughts matter, not their particular expressions in linguistic terms (when we are not in the domain of logic and idealized languages).

Russell also had a deeper interest behind his theory of the definite descriptions. That interest was also epistemological: Russell differentiated between knowledge by acquaintance and knowledge by description, in that the former type of knowledge is the basic one and the latter secondary or just distinct type of knowledge. This doctrine

led him to claim that the linguistic expressions relating to the first type of knowledge are the only really referring expressions ("this", "that", "I", and such). Other expressions by which we talk about entities have at most denoting power, including the ordinary proper names and the definite descriptions (as the abbreviations of the former). This view amounts to an extensive reconsideration of the workings of natural language in the service of knowledge-seeking activities. Similar reconsideration was involved in Frege's perspective. As pointed out, Frege held that thoughts are grasped with the garment of the expressions of natural language, but that the latter tend to be rather defective; they can be ambiguous, for one thing. The primary interest should therefore be focused on the logically "healthy" expressions in the service of finding out the true thoughts. In this respect the empiricism behind Russell's theory of the definite descriptions and Frege's (platonistic) realism with respect to the senses and thoughts led, in effect, to the same result, that of emphasizing the role of logically perspicuous investigations. (Which fact obviously explains in large part why both are considered as the founding fathers of analytic philosophy.)

What about the dissimilarities, or even contraries between the two perspectives? The most important of them lies within the very epistemological and ontological views and their implications. The most striking one is that Frege considered the ordinary proper names, definite descriptions and sentences all to be proper names (*Eigennamen*), thus all of them being referring expressions with *Bedeutungen*. For Russell the logically proper names were the only referring expressions, from which it follows that not any of the three of Frege's subcategories of proper names are referential for Russell. In a sense the situation is neat because the Russellian referring expressions are mainly indexicals and demonstrative expressions. In contrast Frege did not say much about these types of expressions (though there is the often-discussed piece about the I-sense in (Frege 1997b)). The common opinion among the philosophers of language is that Frege's theory fails with respect to the demonstrative and indexical expressions. But the arguments that are offered to support this conclusion all presuppose that the senses and thoughts must be given by linguistic expressions, that all them are linguistically articulatable without a remainder. And that is not true, not even by Frege's own lights (see the paper just mentioned about I-sense). Here I will not go into the issue, but as will become clear later when I develop my own Fregean neurocognitive theory of reference, the semantic equivalence between the senses and the definite descriptions expressing them could not be achieved, but neither need it be achieved: the senses enable proper names to be used to refer without that requiring completely linguistically expressible conceptual contents. Anyway, the main difference between the logically proper names and Frege's conception of proper names derives mainly from Russell's adherence to the Millian account that the sole meaning of a referring expression is its referent.

The first similarity between Frege's and Searle's views is that both allow there to be more than one definite description attached to a certain proper name by a speaker. Frege also allowed that senses differ intersubjectively with respect to one referring expression (as is evident in the earlier mentioned Aristotle note). Judged from Searle's explicit statements it seems not to be clear whether or not he thought likewise. But with what I argued above about the differing weights the descriptions assume in different contexts one can argue further that also Searle allows the intersubjective variance of the descriptions (expressive of the senses of proper names). The obvious reason for this is that in some context some description related to a proper name may

possess zero weight, and some other description, or more, may acquire a weight greater than zero (even if it possessed zero weight in some other context).

In relation to Russell's view the situation is not as clear. It is sometimes held that he maintained "one description per name" view, but at other times the tenor of the discussion seems to be that Russell allowed many descriptions to be associated to a name, and even intrasubjective variance of them. I take it that the second interpretation is correct. My reason is simply that it undeniably fits the textual facts. As seems to be the common view I also assume, with respect to all three theorists considered, that the number of the senses and the descriptions associated to a particular proper name can be left open; nothing crucial hinges on that matter. This follows from what is already said: if the topic or the context affects the senses and the descriptions, it is empirically unlikely that any fixed exact number of them would be a reasonable option to hold.

The second similarity between Frege's and Searle's views is epistemological. Although it is true that nothing very explicit is said about the epistemological features of the cluster theory, we can give voice to a cognitive-cum-epistemic perspective operative behind the scenes. In Searle's theory the descriptions associated with proper names are the guiding factors in the referential uses of those names. And to the extent that proper names have a role in our (mostly non-scientific) dealings with reality, in finding out what reality is like, they are of utmost importance to us. Therefore the senses and the descriptions are the primary cognitive and epistemic factors. Searle does not say this in so many words, but the cognitive perspective is evident in his discussion. For how otherwise than by observing the epistemic relevance could we understand his statements that if the properties related to a (purported) referent did not apply to it, the referent could not be taken to exist. Furthermore the obvious truth is that just as Frege's theory is cognitive because that *is* what the postulation of the senses amounts to in effect, so is Searle's theory because the descriptions assume the role they do. In general it can be maintained that both the Fregean theory and the description theory are cognitive-cum-epistemic theories of reference because they explain the referring power of proper names by other factors than by (the use of) the mere name.

It is arguable, however, that what interest the above comparisons may have it is by itself too general and therefore hides from the view the basic semantic and epistemic differences between Frege and Russell. Strictly speaking Russell was suspicious of the very coherence of the senses (Russell 1905). It is widely acknowledged that this seems to be due to a misunderstanding of the nature of Frege's senses on Russell's part, but the claim still stands: Russell did not countenance senses at all, not for semantic neither for epistemic purposes. However the difference between Frege and Russell can be alleviated easily, and that is what has been done (as the discussion above also shows): the function of the definite descriptions in respect of the *ordinary* proper names is the same as the Fregean senses expressed by the descriptions. They both try to explain the referential role of proper names. But what about Frege and Searle? I think it is evident in that case too that there is no problem; affinities are clear, Searle just focuses on the descriptions. Searle is explicit in stating his aims, viz. to "show in what sense a proper name has a sense" (Searle 1958: 167). It seems to me that if anything it is just this statement that has led to the widely accepted opinion, bolstered most notably by Kripke, that the Fregean sense theory and the Searlean description-cluster theory are in effect equivalent theories, or at any rate that there are not any such

differences between them that would force us to keep them apart as substantially distinct theories of reference. But *that* is not correct as many studies by the neo-Fregeans have shown. And to say it once more, Frege was interested in thoughts; in Searle's paper we do not detect such interest, but only the elaboration of the descriptions as expressive of the senses of proper names. (Or the descriptions identical to the senses - for there seems to be an ambiguity in Searle's paper over this possibility.) However it is Frege's, and by implication Fregean, general cognitive framework that I adopt in this study and not the description theoretic, i.e. the language-restricted view. In short, I take proper names as *condensations* of the thoughts that give rise to the uses of proper names. Another way to conceive of the referential linguistic expressions in relation to the thoughts and senses is to take them as compact reports of the multitude of the neurocognitive elements and processes that make up the thoughts and senses. Of course it does not follow that the descriptions would be irrelevant to the study of reference. Because the descriptions are condensed reports of what is going on inside our skulls within the wet informationally driven biological engines, arguments that take descriptions as their point of departure do hit at least in part also the thoughts and senses. That is, the arguments given by the direct-causal theorists against the description theory, which they take to be equivalent to the senses theory of reference, must be taken at their face value. The reason for us not being able to completely evade those arguments is that we could not yet estimate in any generally viable manner the amounts of information that the descriptions express of the thoughts and senses. That is left for a mature neurocognitive theory to do, towards which I hope I will be able to take some important steps in this work.

Notes

1. Burge's paper still seems to have gone unheeded to a large extent among the direct and causal referentialists. This is especially embarrassing because one of Burge's aims in that paper is to show that Kripkean arguments from proper names being rigid designators have no force against Frege. (Though they may hit Russell's view that proper names are disguised definite descriptions, because that view is overly linguistic.)
2. Nathan Salmon argues that Frege's puzzle about the identity statements is not just a puzzle about identity (Salmon 1986: 12). Other linguistic constructions give rise to the same problems, for example definite descriptions. Thus, claims Salmon, the sentence "Shakespeare wrote *Timon of Athens*" is informative, whereas the sentence "The author of *Timon of Athens* wrote *Timon of Athens*" is not. This observation is correct, but the argument betrays a misunderstanding of Frege's theory. Frege held that in the class of proper names, *Eigennamen*, belong proper names as we understand them currently as well as the definite descriptions. So there is a difference without a distinction in Salmon's argument.
3. Of course the senses and thoughts can become the objects of one's focus, thus amenable to the use of the two extensional principles in their full force. This happens when we talk about what a person believes, as for example when we compare the beliefs of two persons and claim that A's belief that *p* is the same that B's belief that *p*. (Frege apparently held that the "realm of reference" and the "realm of senses" are disjoint ontologically. But I concur with David Bell when he argues that they could not be, i.e. that the senses and thoughts can become objects, hence referents (Bell 1979: 109-10).)

4. It is not relevant for my purposes whether the epistemological view preceded that of the logically proper names or it was the other way around, or that they both might have worked in parallel in Russell's thinking leading to the theory of definite descriptions. If the doctrine of the logically proper names is based on the epistemological view, then one can argue forcefully that it is a bankrupt view. The reason is that since the days of British empiricism and the phenomenalism co-occurrent with it we have advanced. Shortly, there are no primitive self-identifying items in our experience; our experiences and their objects are cognitively mediated. So it is implausible that there are any items we are directly aware of (sense data or whatever). This may be one main reason why the direct-causal theorists have substituted the ordinary proper names for the logically proper names in their arguments.
5. About the technical details of primary and secondary occurrences of descriptions, and the problems with them, and Russell's analysis in general, especially in relation to propositional attitude contexts, see (Linsky 1967: 56-76).

3. THE DIRECT-CAUSAL THEORY OF REFERENCE

Background

The counterattacks to the Fregean sense theory and the description theory emerged during the late 1960s and in the early 1970s. The general theory these arguments brought with them is called, somewhat carelessly, by a variety of names, like "direct theory of reference", "causal theory of reference" and "new theory of reference". Without any carelessness, however, it is better to see these different labels as pinpointing the different tenets the general framework embodies. The theory is new, at least in relation to the sense and the description theories. "Direct" comes from the core claim of the framework according to which proper names do not refer with the help of any mediating Fregean senses expressed by the definite descriptions. Quite the contrary it is held that proper names refer directly to their targets. The attribute "causal" characterizes the mediating link between proper names and their referents. Consequently the notion of directness should not be understood in any very strict literal sense; it is only meant to deny that the involvement of any cognitive factors, like the Fregean senses, would suffice or even be needed for the determination of the referents of proper names. (The absence of *any* mediating factor would be a surd and inexplicable metaphysical quirk, amounting to a magical theory of reference (Putnam 1986a: 5, 46).) In a more general way the direct component of the theory states what the semantic values of proper names are, namely their referents; and the causal component provides the way, at the level of individual speakers and the linguistic community, how the particular referents become the semantic values of their names.

But why I seem to add to the possible confusion with the labels? Why call the theory "direct-causal theory"? The reason is simple. "Direct-causal" captures the thrust of the theory comprehensively, as just described. There is also a substantial reason for my choice of that label. In the actual practice the arguments and points in favour of the theory and against the Fregean sense theory, and the description theory, have become rather meshed. That is only to be expected because there has to be some reference-mediating link between referents and their names, even if the Fregean senses are not accepted, i.e. even if the non-cognitive directness of the reference of proper names is insisted on. Consequently I do not think that my label creates any confusion.

The causal links (connections, chains, networks) are constituted by the situations of referential communication in which proper names are used. The links underlying a particular proper name start in what Saul Kripke has called "baptism", and David Kaplan "dubbing". These events are just name-giving acts like "This is Oscar" or "Let us call that limping brown rat of ours 'Edgar'". As these examples show descriptions can be used in the naming "ceremonies". That is not denied by the direct-causal theory. But the proponents of the direct-causal theory argue that the descriptions used serve at most to *fix* the referent. That is, the descriptions may not even be true of the referent, or they may cease to be true of it without the result that the name in question ceasing to refer to the initially named individual.

I will concentrate on the causal component later, so let me start with the direct component first and review the most important arguments against the sense theory and for the direct theory of reference.

The directness of the reference of proper names

According to the direct component of the direct-causal theory proper names do not refer by the mediation of the Fregean senses, or by any such cognitive factors that uniquely pick out the referent in all cases. Proper names are mere "tags" and nonconnotative appellations. This view has a strong predecessor in John Stuart Mill: "...proper names are not connotative: they denote the individuals who are called by them; but they do not indicate or imply any attributes as belonging to those individuals." (Mill 1961: 20) The contemporary version of that theory was originated mainly by Ruth Barcan Marcus (1961), and after that Saul Kripke added some new notions and argumentative depth (Kripke 1972, 1977). (There has been a somewhat heated debate over the issue of primacy, see (Smith 1995a,b) and (Soames 1995).) According to Marcus: "But to give a thing a proper name is different from giving a unique description...an identifying tag is a proper name of a thing... This tag, a proper name, has no meaning. It is not strongly equatable with any of the singular descriptions of the thing." (Marcus 1961: 309-10) The expression "strongly equatable" I take to express the demand for synonymies and/or analyticities between proper names and the definite descriptions. So proper names, as tags, have no meanings in that strong sense. This leaves it open whether Marcus intends to say that proper names have no meanings that would be relevant to their function as referring signs, or that proper names have no meanings at all. I think that we may safely adopt the first option. At least Nathan Salmon points out that the direct theory of reference does not deny that there could be senses *related* to names. According to his formulation the direct theory only denies that there are no "full-blown Fregean senses" which could serve at the same time all of three referential and cognitive functions of the senses (to be taken up later) (Salmon 1982: 11-14).

When we look at the issue from the perspective of the liberal cluster version of the description theory we get a more coherent picture of the direct component because both Marcus and Kripke (as well as the followers) admit that descriptions have a role in singling out, in *fixing*, the referent of the name in question. Marcus is quite categorical over the point, in fact, for she emphatically maintains that entities are singled out *only* with the help of descriptions, though names differ semantically from descriptions (Marcus 1961: 309). After an entity is singled out and the referent of an expression is determined, the descriptions may not have any referential role to play any more in the later referential uses of the name. Kripke is somewhat ambiguous, for he seems to hold this view also, but then sometimes he says that descriptions are not used in fixing the referent (Kripke 1977: 93-4). Be that as it may, the direct component claims that proper names have no such meanings, or cognitive senses, which would be relevant to their referential functions - not that there are no meanings or senses related to proper names at all.

The direct component of the theory was constructed mainly through counterexamples to the cluster version of the description theory, which Kripke took to be a Fregean theory also. The most precise characterization of the direct component is to be found in Nathan Salmon's book *Reference and Essence* (Salmon 1982). As mentioned the direct theorists take singular terms, especially proper names, to be nonconnotative appellations. This feature of proper names is intended to make evident the comparison to the definite descriptions when these are taken to be expressing the Fregean senses that provide the identifying factors of the referents of particular proper names. In

particular the direct theorists take senses to be concepts, so that the referent of a proper name is the object that uniquely fits the concept that characterizes it (Salmon 1982: 9-10). Moreover Salmon clearly assumes that the orthodox Fregean theory is committed to assimilate singular terms to the model of the cluster theory. As we recall, in this theory the sense of an expression supplies a set of descriptions that the referent best fits and therefore uniquely satisfies (Salmon 1982: 10).

Sense₁, sense₂ and sense₃

At this juncture we need to discuss the partition of the Fregean sense to subcategories, or rather subfunctions. As was pointed out in the first chapter Frege could have characterized the senses in a more detailed manner. Fortunately Michael Dummett took up the task and argued for different ingredients of Frege's notion of sense (Dummett 1973); see also (Dummett 1981: ch.6). But also somewhat ironically the matter advanced in the 1970s in the wake of the direct-causal theory by some of the theory's proponents and writers sympathetic to it. It seems that Tyler Burge was first to introduce the tripartition of Fregean sense on the other side of the theoretical fence (Burge 1977).¹ Salmon also uses Burge's tripartition of the Fregean sense operative in Burge's concise presentation of the counterarguments to the Fregean sense theory of reference (Burge 1977: 356). The purpose of these arguments is to show that there is no full-blown Fregean sense which could simultaneously fulfill *all* the tasks imposed to the Fregean subsenses. It should be noted that the arguments allow that there are, or could be, even all the three subsenses related to proper names without that amounting to the full-blown *unified* notion of the Fregean sense (Salmon 1979: 443). As Burge and Salmon conceive of the issue the first subsense, sense₁ is a purely conceptual presentation of the object. It is what a fully competent speaker associates in a particular way to the proper name (in general, to a referring expression). Sense₁ is a psychological or a conceptual notion. (It should be noted that both Burge and Salmon allow images and such occurments to be included in sense₁, contrary to the categorical denial of their semantic and cognitive relevance of them by Frege.) Sense₂ is the mechanism which *secures* the referent of a proper name and *semantically determines* it: "for singular terms, senses serve as 'routes' to singling out the unique object, if any, denoted by the term." (Burge 1977: 356) So sense₂ is a semantical notion. And sense₃ is the information value of a proper name, the contribution it makes to the information content of the sentences in which it occurs. Sense₃ is a cognitive and an epistemic notion (Salmon 1982: 12). This third subsense explains the puzzles arising in the propositional attitude contexts: being cognitive and epistemic notion sense₃ relates to what a speaker believes or knows. Taking into account the differences in beliefs and knowledge is what any adequate explanation of the puzzles manifested by the propositional attitude contexts requires (that is, the failures of the substitutivity of identicals and existential generalization). A concise formulation of the Fregean sense theory, which unifies the three subsenses, is that proper names have senses that are "simultaneously the conceptual content, the semantical method of determining reference, and the cognitive content, all at once." (Salmon 1982: 23)

Frege held that the sense is a uniform notion in its functioning as a cognitive entity in providing the *Erkenntniswert* for the referential expressions (as well as thoughts made up of sense elements). Consequently he did not make any very reflective differentiations within that notion. But it seems to me that the subsenses can be defended by comparing them with some statements by which Frege characterized the

notion of sense. Sense₁ most likely corresponds to the sense as the mode of presentation of the referent. In fact that is Burge's view (Burge 1977: 356). Sense₂ is a more problematical for it would seem that the way of presentation of the referent is also what secures the referent, especially when it is demanded that the presentation is uniquely identifying. As an argument for the differentiation between sense₁ and sense₂ Burge notes that often a complete account of the mode of presentation may be insufficient to determine a unique object. The reason is that the individuation may also depend on nonconceptual but contextual relations to the object (Burge 1977: 357-8). As a general observation this point is well taken, and it is also widely assumed as we will later see. But the assumption on which it rests and which is also widely accepted is quite problematical. The assumption is that for Frege the senses were thoroughly conceptual in being, indeed, non-contextual. I will argue in sections 4.3.3 to 4.3.5 that this is not so at all. Sense₃ corresponds to Frege's notion of indirect *Bedeutung* (customary sense) in propositional and *oratio obliqua* contexts. Burge's argument against the identification of sense₃ with sense₁ is that the descriptions that express the senses of proper names for different persons may differ from one another to the extent that there is no shared sense to function as the object of their respective beliefs (with respect to the proper name in question) (Burge 1977: 359). However this interpretation of the shared indirect referent, i.e. the customary sense, may not be required of the senses. Frege allowed intersubjective variation of the senses related to proper names as is evident in the famous Aristotle note; as long as the referent is the same it does not matter if the respective senses of the speakers are not the same.

In this connection it is in order to note again that the characterization of sense₁ as psychological or conceptual is not wholly correct, if it is intended to remain faithful to Frege's notion of sense. As is well-known Frege denied senses any psychological aspects and relevance. Grasping the senses and thoughts is a psychological act, but in this act we do not as it were create or construct the senses, we only get in cognitive relation to them. Moreover it also seems to me that it is not an altogether happy move to take sense₁ to be a conceptual notion, if one intends this to conform to Frege's notion of concept (*Begriff*). The reason is that in his logic and semantics Frege made a distinction between objects and concepts (functions in general). The senses can be taken to be objects also, as I mentioned earlier. For example they are capable of assuming the role of indirect *Bedeutungen* as customary senses. But here we may give the opponents the benefit of doubt and allow that the senses are only conceptual. Still we can argue, in conformity to my proposal, that this conceptual interpretation of sense₁ can be justified on the Fregean grounds. Take for instance the sentences "the pupil of Plato and the teacher of Alexander the Great". Now "the pupil of Plato" and "the teacher of Alexander the Great", which express concepts (*the pupil of Plato* and *the teacher of the Alexander the Great*) and even the functional expressions "pupil of _" and "teacher of _" can be taken as concept expressions in the sense that their *Bedeutungen* are not objects. They are in need of arguments to "saturate" them, i.e. to make them expressions of whole truth-evaluable thoughts. I think that this is in line with Frege's logical semantics, so the conceptual interpretation can be sustained, at least for the sake of the direct-causal theorists' arguments. However the senses of proper names need not be *purely* conceptual in the sense of being purely qualitative. This is easily seen for they contain parts having other individuals as their referents ("the pupil of *Plato*" and "the teacher of *Alexander the Great*"). But the argument for this will be postponed till chapter 4.

Salmon adopts the Burgean taxonomy of the subsenses, as i) the mode of presentation of the referent, ii) the determiner of the referent as the mechanism by which the referent of a proper name is secured and iii) the information value (or the object of thought) of an expression. With this taxonomy the core claim of the direct component receives an accurate formulation. As a consequence the direct-causal theory does not deny that the senses are *associated* with singular terms. What it denies is that there is such a thing, the Fregean sense proper, which fulfills all the three functions of the subsenses (simultaneously). In particular the direct component can allow that proper names have associated with them any of the three subsenses (or even "the ordered triple consisting of all three" (Salmon 1982: 14, 23)). It is evident in many writings of the direct-causal theorists that they take the Fregean theory to be a theory that assimilates proper names to the definite descriptions and makes the former descriptive through and through, the function of which is to express the properties the referent satisfies. For example still recently Scott Soames has maintained that, according to the sense theory, speakers associate different descriptive contents to proper names they use. Soames elaborates on that phenomenon (Soames 2002: 72-86), but not surprisingly, as the direct-causal theorists as Millians are wont to do, he denies that the association of the descriptive contents has anything to do with the *semantic contents* of proper names. Those contents are solely the referents of those names (or the singular propositions expressed). When this view is combined with the denial of the existence of the full blown Fregean senses, the direct component is given a positive edge: certain singular terms are entirely non-descriptive. The way they refer is by an assignment of a value (like one does with free variables) and by not satisfying any conceptual condition given by the descriptions (Salmon 1982: 16, 22).

Indeed, Salmon claims that proper names are "thoroughly descriptive" according to the "orthodox Fregean" account. This means that they do not include any constituents that indicate a further individual. Thus only thoroughly descriptive terms express purely conceptual senses, and the properties involved are supposed to be purely qualitative and/or general in that they do not involve a direct reference to any other individual (Salmon 1982: 17-20). For example the sentences "the wife of Socrates" and "the mother of Jesus" are not thoroughly descriptive because they both include a name of an individual. (Though these examples are not very convincing at all, for presumably the senses they express can be given by the descriptions "a wife of the most famous ancient philosopher who drank hemlock as a punishment for philosophically corrupting people" and respectively "the mother of the savior of the humankind".) As Salmon says explicitly "...the orthodox theory of reference claims that names refer to someone or something entirely by way of associated general properties leaving no room in its account for either the linguistic or the extralinguistic context." (Salmon 1982: 32) and that Frege "seems to have held a particularly strong version of this theory...it seems that he held that all proper names are thoroughly descriptive." (Salmon 1982: 21) (see also (Salmon 1979: 444-5)). Here we should note that in this latter quotation Salmon seems to hesitate in his categorical attribution of the view to Frege, but a few pages later he has made the transition and made his mind up firmly, for he generalizes and claims without any hesitation that "Fregean theories involve the strong thesis that singular terms are not only descriptive, but thoroughly so." (Salmon 1982: 23) Consequently it is this interpretation of Frege's theory of the sense and reference that we will assume to be the target of the arguments given by the direct-causal theorists. For if the weaker interpretation is accepted it is left open that proper names might not be "thoroughly descriptive" after all. (Which

is precisely what the neo-Fregeans have argued time and again, and which will be newly argued in sections 4.3.4 and 4.3.5.) But before we deal with all these arguments we must discuss one crucial notion that the direct-causal theory has brought with it.

Rigid designation

Kripke introduced the notion of *rigid designator* (Kripke 1972) (see also (Stanley 1997: 555-6)). A proper name is a rigid designator if it refers to the same individual in every possible world where that individual exists. It is to be understood that the actual world is what I call the "primary referential source": the proper name in question refers to the same individual in every possible world to which the name refers in the actual world. Hence "Josif Stalin" refers to Stalin, the dictator of the Soviet Union, in all the possible worlds where he exists, even if his name in some, or all of them, is different in those worlds – and even if he were someone totally different individual by his deeds, like an Ukrainian peasant near Kiev.² Kripke's favourite type of argument for the rigidity of proper names goes along the following lines. Take a description (or a set of descriptions). Assume that it refers to someone or something in the actual world via the proper name to which it is associated. Now take another possible world or a counterfactual situation, distinct from the actual, and inquire to whom or to what the description refers there. As an example let us use the sentence "the dictator of the Soviet Union could have been shot during 1930s". The referent of the description could be different from the one it is in the actual world, but that the proper name, "Josif Stalin", associated with the description still refers to the same (actual) individual, Stalin. (Or that the actual Stalin exists there but has some other name.) Hence "the dictator of the Soviet Union" may refer to Trotsky in some possible world, or even to Ordzhonikidze in some other one. But "Stalin", or whatever his name is in the possible world, would still pick up the actual Stalin in that world (if he exists in it). Notice, by the same token, that in this example it is to be understood that "Trotski" and "Ordzhonikidze" also refer to actual Trotsky and Ordzhonikidze. So in a way it might seem that Kripke's setting of his argument already presupposes the notion of rigid designator he argues for. But that does not follow, for all that is required is that the description does not necessarily refer to (actual) Stalin in those other possible worlds - as it does not seem to, intuitively.

The above example is a modal one in that it has to do explicitly with alethic modalities concerning actual world individuals. But modal, or intensional, contexts in general are not required for the argument from rigidity. Non-modal sentences, or what Kripke calls "simple sentences", will do as well. To give an instance of them take the sentences "Frege had moustache" and "the most renowned logician of the 19th century had moustache". The first sentence is true with respect to a world *w*, if and only if, Frege (himself) had moustache in that world. The second sentence is true in a world if the person who is the most renowned logician of the 19th century in that world had moustache; and that person need not be Frege himself (although he might have been called "Frege"). This short example reveals that the point of rigidity, as Kripke has emphasized, has to do with the truth-conditions of the sentences with singular terms, not, as is sometimes thought, with the distinctions among the scopes the rigid designators and the descriptions assume in the modal contexts.

As said the important semantical feature of the rigid designators is that, where and when a rigid designator does refer at all, it refers to the actual individual, the

individual *we* use it to refer to. Kripke states that it is to put the cart before the horse if one assumes that we must antecedently make sense of the criteria of transworld identity of individuals, rather than being able to talk about them, to refer to them, quite unproblematically as the very individuals they are actually (Kripke 1972: 270). It is because we refer by proper names that are rigid designators that any transworld identity criteria are spurious: the referents remain the same actual individuals. The actuality of the individual and the focus on the truth-conditions is made clear also in another place where Kripke says that "...we begin with the objects, which we *have*, and can identify, in the actual world. We can then ask whether certain things might have been true of the objects." (Kripke 1972: 273)

One central consequence of the direct referentiality and the rigidity of the singular terms is that identity statements, both trivial and informative, are necessary if true. When "a" and "b" refer to the very same entity, and are both rigid designators, they refer to that very same entity in all the possible worlds where that entity exists. Hence "a is b" is necessary if true. But that a is b is not known *a priori* (see the epistemological argument below). So "a is b" is *a posteriori* truth also. To my mind Quentin Smith has argued convincingly that Marcus already had the epistemological argument and in effect the notion of necessary *a posteriori* identities (Smith 1995b: 232-4). Smith also explains the relevant metalevel-object level distinction involved. When "a" and "b" have the same referent and are directly referential rigid designators, then "a is b" is *a posteriori* identity. That is a metalevel fact, for the referents of "a" and "b" could have different names than they do have (i.e. "a" and "b" here). The necessity of the statement is an object level fact: that a is b (or that a is a, or b is b, for that matter) is not an empirical fact because an entity's being identical to itself is *a priori* fact and true in every world.

One more point about rigidity. It has sometimes been thought that definite descriptions could not be rigid. But it is commonly noted that some definite descriptions are rigid designators. For example the description "the only even prime number" is such. The reason is of course that 2 is the only even prime number, i.e. the description is true in all (mathematically accessible) possible worlds. In Kripke's terminology this is *de facto* rigidity in that the referent happens (as a mathematical necessity in this case) to be the self-same in all possible worlds. But proper names are rigid *de jure*: the semantical rules related to proper names and manifested in their referential uses account for the fact that any proper name refers to the same individual in all possible worlds irrespective of any descriptonal content related to it. Or so the direct component of the direct-causal theory also states.

What is the relationship between rigid designation and direct reference, or direct referentiality? If directly referential term is a term which refers without the mediation of the Fregean senses, or an associated cluster of descriptions (expressing the properties of the referent), and it contributes to the singular propositions only the object it is the name of, there is not a sufficient distinction between directness and rigidity. Francois Recanati argues that rigid designation is not the same notion as (type) referentiality (Recanati 1988). The main reason is that it is not necessary for a referential term that its referent exists. Recanati goes on to draw a distinction between rigidity and direct referentiality with the help of the notion of singular truth-conditions. A directly referential term is one that indicates that the truth-condition of the utterance S(t), in which the very term occurs, is singular. Truth-condition is singular if and only

if there is an object x such that the utterance is true if and only if x satisfies $S()$, i.e. what is attributed to x (Recanati 1988: 113-4). As Recanati says a referential term is such that it indicates its own rigidity. He postulates that the means by which this condition is achieved is a general feature "REF", a part of the linguistic meaning of directly referential terms (Recanati 1988: 114-5). But note that the feature "REF" is not intended to be any kind of conceptual component of the sense of any proper name, even if it functions as a part of the *linguistic* meaning of a term (the linguistic mode of presentation of the referent in Recanati's framework). The semantical force of "REF" is that it indicates that the term is directly referential (and rigid), and this has the consequence that the important property of directly referential terms, especially proper names, so to say their referential nonconceptuality, is preserved here.

Recanati's explication of direct referentiality is "in principle" type argumentation; he says in fact that he is *defining* the notion of referentiality. This follows from the fact that although some definite descriptions are rigid *de facto*, in general the definite descriptions are usually not taken as being referential devices in their (Donnellanian) attributive use. But it seems to me that for my argumentative purposes I can keep the looser characterization of direct reference as reference that is not mediated by any full blown Fregean senses, with the result that proper names as rigid designators are directly referential expressions, unlike (most non actualized) definite descriptions which can be taken to express (at least parts of) the Fregean senses. (On this see section 4.5.)

The main arguments against the Fregean sense and the description theories of reference

The modal, epistemological and semantical argument

Kripke used three main arguments against the Fregean sense-cum-description theory, as he understood it at the time. These versions are aptly summarized by Salmon, both in the book mentioned above and in his review article (Salmon 1989). The three arguments are i) modal, ii) epistemological and iii) semantical. Let us take a short look at them in turn. It should be noted that the focus of them all is on Fregean assimilation of sense₁ and sense₂; sense₃ is left out of consideration for the time being. Argumentatively this is straightforward, for if sense₁ and sense₂ can not be assimilated to each other, there are no full blown unified Fregean senses, consequently no need to be concerned with sense₃ (however see (Salmon 1979: 449-50)).

The modal argument is based on the presupposition that proper names and the definite descriptions are synonymous with one another and the descriptions are definitionally true of the bearers of proper names. Thus according to Salmon statements like

"Shakespeare, if he exists, wrote '*Hamlet*', '*Macbeth*', and '*Romeo and Juliet*', then he is Shakespeare."

and

"If anyone is an English playwright who is the sole author of '*Hamlet*', '*Macbeth*', and '*Romeo and Juliet*', then he is Shakespeare." (Salmon 1989: 446)

express logical truths and are analytic sentences and so true in every possible world when the descriptions are substituted for proper names. (Obviously the descriptions are to be understood as expressions of the corresponding senses, as mentioned above. As Salmon says the "wrote" part is taken to consist of any and all descriptions purported to be about Shakespeare.) But it is easy to imagine a possible world in which Shakespeare, the actual man, did something else than write the plays. So the statements are not true in every possible world, hence the description theory is refuted because the synonymy between the proper name "Shakespeare" and the descriptions fails: synonymy does not effect a change in the truth-values.

The epistemological argument is a close kin to the modal one. Assuming still with the direct-causal theorists that the statements above should be analytic, they should also be knowable *a priori*, based only on the concepts involved (or expressed). But again with only a little help from imagination and circumstances will emerge in where Shakespeare did not write the plays. Since this possibility is not ruled out by the contents of the concepts involved, it follows that the respective statements are known only *a posteriori*.

The demand of the modal argument that there must be synonymies between the descriptions and proper names can be readily understood when it is realized that the direct-causal theorists have taken the Fregean sense theory to be a theory of the *meanings* of proper names. (It seems that this is inferred from the label "sense" in the sense theory. But more of this in section 4.1.) But where does the demand of *a priori* knowledge come from? The only answer I have been able to fathom from the direct-causal literature is that once the equivalence of Russell's theory of the definite descriptions and Frege's theory of sense and reference has become assumed, it has been easy to import the knowledge demand from Russell's side. In his theory there could not arise any doubts as to whether or not two expressions are co-referential if they are logically proper names. Now if one loosens the search for the logically proper names and turn to the ordinary proper names, as the direct-causal theorists have done, the epistemological argument follows naturally. (Jason Stanley has also lately stated that according to the description theory of reference we possess *a priori* uniquely identifying knowledge about the referents of proper names (Stanley 1997: 565). But he gives no argument for this statement, and it is hard to see what argument would *force* upon the description theory such demand. What comes to Frege, he denied any such demand (see section 4.1).)

The semantical argument is different in type from the first two. It derives from Keith Donnellan's paper "Proper names and Identifying Descriptions" (Donnellan 1972). Take the description "the Greek philosopher who held that all is water" and equate it with the name "Thales". Suppose, however, that the man from whom the use of "Thales" originated never held such a view, but that there was other fellow philosopher who put forth such wet metaphysics. Now it is natural to ask to whom does "Thales" refer, and according to Salmon this is a clear semantic question with a clear semantic answer: the name would refer to the first individual, but not to the water metaphysician. The semantic question is supposed to probe the notion of reference *simpliciter* (Salmon 1989: 449). (Of course this argument presupposes that our semantical and referential "intuitions" are uniformly valid, otherwise the talk of reference *simpliciter* does not make much sense anyway.)

The first two arguments are the same in type as we saw Kripke's general argumentation for the rigidity to be. They are designed to show that the (Fregean) description theory fails if it is supposed that proper names are synonymous (mean the same, have the same senses) as the associated descriptions. This view of the relationship between proper names and the definite descriptions is evident in Kripke's essay *Naming and Necessity*. He states there quite a few times that the description theory is best understood as the theory of *meaning* of proper names. This involves that proper names are analyzed by the descriptions to give synonymies (and thereby definitions) of the former (Kripke 1972: 277, 281). (I use the 1972 version of Kripke together with the 1980 version rather indiscriminately. The reason is that there are no substantial differences between them, except for the Preface in the 1980 version.)

As was pointed out above, Kripke invokes the notion of fixing the referent, or determining the referent - as did Marcus earlier (Kripke 1972: 258-9). He admits that the description theory can be understood in that sense also. A referent can be fixed by a description, or descriptions. But it must be allowed that the description(s) used is (are) not true of the referent, or that it (they) might become false of it later.³ One may fix the original referent of "Thales" by giving a description that says that he put forth a watery metaphysics, but as the semantic argument purports to show, this description might not enable us to refer to the original individual named Thales because it is false about him and we may not know anything nontrivial about him. But, according to the claim behind the semantic argument, the name *does* refer to the original Thales in spite of that.

Salmon takes the semantical argument to be the strongest one both against the description theory and *for* the primary thesis of the direct component (Salmon 1982: 29). He also claims that the semantic argument reveals that contextual factors are involved in the constitution of sense₂, i.e. the way the referent is determined. I take it that that is because of some reference conducting features in the naming situation, and not the description, that the original Thales remains the referent. Moreover that happens in such a way that it becomes clear that sense₂ is not a purely conceptual entity. He finds the moral of the semantic argument to be that it shows that the surrounding settings of the speakers are crucial in determining the referents of proper names. In particular this contextual involvement "is true not only of the extralinguistic setting, in which the referent is to be found, but also of the linguistic setting in which the term is used or was learned by the speaker, i.e. the history of the use of the name leading up to the speaker's acquisition of it." (Salmon 1982: 31)

Salmon's diagnosis of the modal argument is that its force comes from an underlying intuition according to which a proper name continues to refer to the same person or object in counterfactual situations, even if he or she or it has not a single one of the properties by which we actually identify him or her or it (Salmon 1982: 26, 32). (I take it that this does not concern about the essential properties of an individual, if such there were, and if such are used in the descriptions.) This intuition, or at least something very similar to it, lies also behind the epistemological argument. But what about the semantical argument: what kind of evidence can it claim to possess as support? Remember that it is taken by Salmon to be the most persuasive of the three arguments, so presumably it is reasonable also to assume that the evidential base it has should be stronger than those of the first two, something stronger than a mere modal intuition.⁴ But, again, let us postpone the treatment of these arguments and their

evidential bases to the next chapter. There are two further arguments that Salmon gives in another occasion against the Fregean theory, and by implication against the (Fregean-cum-)description theory.

Twin Earth and content variance argument

In addition to the three previous arguments there are two other types of argument that are quite frequently given against the Fregean theory. In his book *Frege's Puzzle* Salmon formulates them concisely ((Salmon 1986: 66-8; but see also (Salmon 1989: 448-9)). The first is a variant of the Putnamian Twin Earth argument and purports to show that the purely conceptual content is not the information value of a proper name. However there is one crucial presupposition behind the Twin Earth scenario. It is that one's internal psychological or cognitive state determines the senses one is grasping (to put it in Fregean terms) or the concepts one is entertaining (to put it in modern terms). From this it follows that when the psychological or cognitive states are the same, the grasped or entertained concepts are the same. Now suppose that Hubert's wife has a physically identical *Doppelgänger* in Twin Earth. Hubert himself weighs 165 pounds whereas his *Doppelgänger* weighs 165.000000001 pounds. The wives of both Huberts assert that "Hubert weighs exactly 165 pounds". As the wives are exact replicas of each other, their respective psychological or cognitive states are also the same (by hypothesis). Therefore the conceptual contents of the states are the same. But the information encoded by the two respective statements is different. The statement of the Earth wife about Hubert is true, but the statement of the Twin Earth wife about Twin Hubert is false. It follows that the information values of the two names "Hubert" differ from each other. Therefore the conceptual content of a proper name is not (uniformly) the same as its information value. This argument is often put in terms that the Twin Earth examples confirm the case against the Fregean view that sense determines reference (although the case is usually given with respect to the natural kind terms). The reason is evident: the contents interpreted as the senses of the two wives are the same but the referents are not.

Salmon's second argument I call, for the want of a better label, "the content variance argument". It is directed against the proposal that conceptual content is one ingredient of the information value of a statement, the other ingredient being the referent. The argument receives its rationale from the subjective aspects of the conceptual contents of proper names. Conceptual contents vary from person to person because different persons have different conceptions of the bearer of the proper name in question. You may think of Aristotle primarily as the teacher of Alexander the Great, and I may think of him primarily as the most famous pupil of Plato. Moreover the way the conception is acquired differs from a person to another: some may be well acquainted with the referent of the name, some others may have only heard of some of his features, others still may have seen him but forgotten about that, and so forth. In general from these kinds of facts it follows that the information value of a proper name, or a sentence containing that proper name, will vary from person to person because, if conceptual content is a part of it, the content varies from person to person. But this "clashes sharply with the original, natural idea of a sentence-for example 'Socrates is wise'-encoding a single piece of information (the information that Socrates is wise). The sentence...encodes the same information for you as for me...it does not encode a piece of information *for* someone." (Salmon 1986: 68). The reason is, according to Salmon,

that the relation of encoding information between sentences and information itself is as objective as the semantic attributes of truth and falsehood.

Salmon points out two caveats to the content variance argument. The first requires that the encoding relation must be *relativized* to a particular type of use of the sentence. In other words the proper name the sentence includes is understood to refer to an individuated person or object. "Aristotle wrote *Metaphysics*" encodes the true information about Aristotle, the Greek philosopher, and false information with respect to the Greek shipping magnate Aristotle Onassis, or the 15th century master builder Rodolfo "Aristotle" di Fioravanti. As Salmon correctly points out it is necessary to relativize the encoding relation to the types of use to achieve nonambiguous references for proper names to begin with. However this relativization is not the same as the relativization to a particular conceptual content because on Salmon's view the conceptual associations are not relevant to the semantic attribution and it is the latter what the relativization of use is all about. After the relativization to the types of uses of a proper name the encoded information is the same for everyone and with a single truth-value. Salmon's second caveat is that, once the relativization has been made, there are still left the different conceptual contents that the different speakers possess. But here Salmon invokes the distinction between information that is semantically encoded and information that is (only) pragmatically imparted. Typically a sentence, or an utterance, imparts more information than it semantically encodes. "Aristotle wrote *Metaphysics*" encodes the information, roughly, that Aristotle, a Greek philosopher, wrote the book that is called "*Metaphysics*". But it imparts pieces of information that depend, in a case of a particular speaker, on the conceptual associations of the pieces of knowledge he possesses, for example that the author of *Metaphysics* was a prolific writer and lecturer, and did not like Plato's theory of Forms.

Kripke's semantical arguments from ignorance and errors

Kripke has also given arguments with a strong empirical bent in addition to the modal and epistemological ones against the Fregean description theory. Those arguments were devised to show that proper names and the definite descriptions do not have exclusively equivalent modal profiles therefore they could not have the same senses, or meanings. That leaves open the option of the Fregean description theory being a theory of *reference*. Kripke's empirical (-cum-intuitive) argument from the causal and communicative links (beginning from the fixings of the referent) is designed to show that the "Fregean description theory" fails even in that case. So while the definite descriptions do not provide synonymies or other such strong semantic equivalences for proper names, it is allowed by the direct-causal theorists that the referents could be fixed with the help of the descriptions (expressing senses) and that these are associated with proper names, it is denied that they could secure the referents in the long run; only the causal links are able to do that.

Kripke is explicit in saying that the arguments are directed against the *Searlean* cluster version of the theory – but there is certain, so to say, explicit undecidedness in Kripke's overall view, especially in the 1980 book edition, as we will see. So it is in order to give here the Kripke's characterization of the cluster version because that is his official explication of the theory he attacks (Kripke 1972: 280-2, 285; 1980: 71).

- (1) To every name or designating expression X , there corresponds a cluster of properties, namely the family of those properties Φ such that A believes ' ΦX '.
 - (2) One of the properties, or some conjointly, are believed by A to pick out some individual uniquely.
 - (3) If most, or a weighted most, of the Φ s are satisfied by one unique object y , then y is the referent of ' X '.
 - (4) If the vote yields no unique object, ' X ' does not refer.
 - (5) The statement, 'If X exists, then X has most of the Φ 's' is known *a priori* by the speaker.
 - (6) The statement, 'If X exists, then X has most of the Φ 's' expresses a necessary truth (in the idiolect of the speaker).
- (C) For any successful theory, the account must not be circular. The properties which are used in the vote must not themselves involve the notion of reference in such a way that it is ultimately impossible to eliminate.

The points (1) and (2) make it evident that Kripke takes the cluster version being a theory for an individual speaker reference. (3) states the core claim of the cluster version according to which the descriptions single out a unique referent, i.e. the individual that the descriptions apply to or are believed to be true of. (4) just adds the corollary that the descriptions must single out only one referent. (5) and (6) spell out the (alleged) epistemic and metaphysical relations between proper names and the definite descriptions, namely apriority and necessity (on which the modal and epistemological arguments focus on). And finally the condition (C) can be taken to express an obvious and minimal, demand that *any* proper theory of reference must satisfy.

It must be noted already here that the inclusion of (5) and (6) to the cluster version is a very debatable move. Without any specific arguments the cluster version, let alone the Fregean, should not be burdened with the strong claim that a speaker *knows* that there is (or even should be on the theory's count) any *a priori* or necessary connections between proper names and the definite descriptions related to the former. I think that this demand simply betrays a misconception on Kripke's part of the intents of the Fregean description theory of reference - even if we ignore the fact that the notions of apriority and necessity, as well as their close relatives, synonymy and analyticity, are very problematical notions indeed. At least I do not know any place where the cluster theorists, and description theorists in general, have voiced any such strong demands to be fulfilled by their theory. Consequently I will drop the desiderata (5) and (6) from now on.⁵

Kripke's first argument takes an issue with the desideratum (2). According to (2) a speaker must have some beliefs as to what properties the referent possesses such that individuate it as the unique referent of the proper name in question. Kripke's argument purports to show that these kinds of beliefs are not required for the speaker to be able to refer to the individual with that proper name. Take for instance the name "Feynman", as used by a person who does not know anything else about Richard Feynman than that he was a famous physicist. Still the speaker "uses the name 'Feynman' as a name for Feynman." (Kripke 1972: 292). (Of course this assertion has to be understood as saying that the speaker uses the name "Feynman" to *refer* to

Feynman, and not only that “Feynman” belongs to the linguistic category of proper names.) There is also the danger of circularity lurking behind the innocent looking desideratum (2). One may be able to refer to Cicero only by the description “Cicero denounced Catilina”. But this description involves another name that has to be accounted for. Because this piece of information is all that the speaker may possess, it is easy to propose that the reference to Catilina goes via the description “a man denounced by Cicero”, the result being referential vicious circle. Another version of this difficulty given by Kripke concerns the name “Einstein”. One may describe the referent of that name as the discoverer of the theory of relativity, but be completely incompetent to tell anything about the theory to back up his reference by the description, except that it is Einstein’s theory.

These arguments purport to show that knowledge or beliefs – that is, Fregean thoughts and the definite descriptions expressing them - about the bearer of a proper name are not necessary for the reference of that name. Kripke offers also examples by which he aims to show that knowledge or beliefs are not even sufficient for a successful reference with a proper name. The argument these examples are designed to support is directed against the desideratum (3). In short, let us assume that some person named “Schmidt” proved the incompleteness of arithmetics, but that due to some mishap Gödel (that is, Kurt Gödel, the renowned logician with paranoia) got hold of that proof and has since been taken as the man who proved the incompleteness of arithmetics. Now Kripke states that it is clear to him that despite the malicious circumstances the name “Gödel” is not used to refer to Schmidt; it is used to refer to Gödel (Kurt Gödel) himself (Kripke 1972: 294). Also Einstein comes back to back up Kripke with respect to the insufficiency argument. For someone who believes that Einstein (*the* Albert Einstein) invented the atom bomb is, according to the description theory, not referring to anyone in particular, for no one invented the bomb single-handedly - or at least is not referring to Einstein (for he did not invent the bomb, he merely derived the formula which later played a part in the revolution of the elementary particle physics, with the well known ramifications; and he signed the famous letter to Roosevelt recommending that A-bomb should be produced). But to claim that that use of “Einstein” does not refer to Einstein is according to Kripke “... simply...false.” (Kripke 1972: 295)

So these are the main types of arguments that the direct-causal theorists have given to show that the sense and the description theories are false to the core, and that proper names are directly referential expressions. As said above, the directness has to be understood in a loose sense: references of proper names *are* mediated, but not with such cognitive factors like the Fregean senses expressed by the definite descriptions that would function as the referent determiners once and for all. Reference of proper names is mediated instead by causal-cum-communicative relations. It is to this component of the direct-causal theory that we now turn.

The causality of the reference of proper names

Kripke’s picture of reference

Kripke sketches his “picture” of causal reference basing it on his arguments purporting to show that descriptions (expressing speaker’s beliefs about the properties of the bearer of a proper name) are neither necessary nor sufficient for successful reference

with the related proper name (Kripke 1972: 298-300). I shall first sketch Kripke's sketch. Then I will review the improvements Michael Devitt has proposed in investigating the communicative and causal factors that are central to the causal component. I hope that this gives us sufficiently comprehensive account of both the developmental course and the substance of the direct-causal theory in order for me to be able to engage the critical stage of this work.

Kripke's picture of the causal chains of reference begins with naming situations ("baptisms") where for instance a baby is given a proper name by the parents or caretakers. (Baptisms should be taken in a loose sense: a name is a name of any kind of an individual; it can be given even unintentionally as long as it sticks; also nicknames are included.) Those who have given the name are the primary users, at least before they have told anyone else what the name of the baby is. To whom they tell the name along with other things, like "Paul babbles a lot", catch on with the name and can then use it to talk about the little Paul to other people. So the causal chain is started and it is weaved on from one occasion to another when the name "Paul" is used (with descriptively given information in most situations). That is why the ignorant speaker will eventually be able to use the name "Paul" as the name of Paul Joseph Goebbels, even if he knows only that Goebbels was a German propagandist. In the same manner the innocent speaker can refer to Richard Feynman the physicist, because he is in the receiving end of the causal chains from the baptism of that Richard Feynman. The speaker does not have to remember from whom or where or when he got the name: the causal chains take care of the reference to a unique individual. It is easy to see that this is only a general characterization of the causal reference relation, indeed more of a picture than a theory. What Kripke specifically means by his talk of picture is that he is not offering any philosophical theory of reference that should be given by stating necessary and sufficient conditions for reference (excepting the minimal condition (C)), conditions which would be immune to all counterexamples. This is evident when he points out that his picture would need to be given further conditions to be satisfied because "of course not every sort of causal chain reaching from me to a certain man will do for me to make a reference" (Kripke 1972: 300).

The causal referential chains in question are mainly constituted by communicative situations in which the name is used. They start in "baptisms", and descriptions can be used ("let us call this limping brown rat of ours 'Edgar'"). After the baptisms everyone who uses the given name has inherited it via the causal links from the baptism event to the current uses. Moreover it is claimed that the current user *intends* to continue to use the name the same way, to refer to the same entity that the previous user did (from whom he caught the name). It is common among the direct-causal theorists (and the Millians in general (Soames 2002)) that these intentions suffice to retain the original referent. For otherwise the causal links may get meshed and changed without us even noticing it: the hearer may think that the name the speaker uses would be a perfect name for his pet aardvark instead and consequently continue to use it so breeding referential confusion. Kripke also realizes that because of the need for the co-referential intentions his picture of reference will not do as a theory (according to his stringent conception of a philosophical theory), for were he to offer it as such, it would violate his own condition (C). (Unless, of course, the notion of intention to *co-refer* could be eliminated, or explained by other notions not invoking further referential notions.)

Devitt's theory

Michael Devitt has made a quite thorough work in providing many of the missing pieces of the causal theory (Devitt 1981). Let us take a look at his version of the theory and see how it improves on Kripke's rather plain picture. As we saw the first version of the direct-causal theory of reference were inspired by modal considerations (related to the notion of rigid designator). This is understandable because Marcus and Kripke had been doing some seminal work in modal logic earlier. But much more empirical and naturalistic versions of the theory has been put forth since and Devitt's theory is the most thoroughly developed of them. Devitt is also explicit with respect to his background assumptions. He takes language to be a natural phenomenon without giving to it any privileged place in the scheme of things. Consequently reference has to be explained on naturalistic grounds using the best tools that the related scientific disciplines provide. (In this respect I share Devitt's stance wholeheartedly.) However Devitt does not endorse the modal and epistemological arguments because he is suspicious of the explanatory value of (the notion of) possible worlds (Devitt 1981: 212-3; Devitt & Sterelny 1987: 31-2). This leads him naturally to emphasize empirical arguments against the description theory and in favour of the direct-causal theory.⁶ (But also Devitt takes the description theory to be Fregean.)

Devitt gives three arguments against the description theory (Devitt 1981: 14-23). They are only slightly different from the ones that Kripke have given, but for comprehensiveness' sake I will present them here. According to Devitt there are cases in which the sense theoretic reference fails because i) we do not or can not associate an appropriate description, expressive of the sense, with a name, ii) where the description identifies a wrong object, and iii) the description identifies no object, although the name is not an empty one (i.e. it does refer). As to argument of type i) Devitt states that it is no use to offer descriptions if they contain information (or names) which needs to be referentially specified in their turn. As an example Devitt also gives the description "Einstein = the discoverer of the theory of relativity" of which he claims that one does not make a referential use with it if one could not tell independently of the very description what the theory of relativity is. What about the cases in which a description (or descriptions) pick(s) out a wrong object? Devitt claims that people usually have many false beliefs about various objects and persons. But whether this really is so seems to me very doubtful: it seems to me that *usually* people are not that ignorant. (Do people really *usually* have *many* false beliefs?) And how false the beliefs have to be to bring about a referential failure? But let us let these problems pass for the sake of the argument. Then it is an empirical fact that people at least sometimes do identify the wrong object by their beliefs (expressed by descriptions) but the names they use seem to refer to a definite object To illustrate the third type of argument, when a proper name is not an empty one but no description is true of its referent, Devitt discusses the case of Jonah from the Bible. Think about the biblical descriptions of Jonah. Most likely all the non-trivial of them are wrong, or we may suppose so (being eaten by a whale and surviving?!). Still there may have existed a man named "Jonah", but all we have at our disposal are the false descriptions about him. The traditional description theory predicts that in this case no one has referred to that Jonah, at least not before the hoax was discovered and it was (supposedly) also discovered that there was a man about whom the biblical stories are told. But that result is absurd, claims Devitt. For surely the name "Jonah", as it has been used

through centuries, has always referred to that Jonah about, and around, who the legend was weaved.

The causal component in Devitt's theory is very much the same as Kripke's, but Devitt spends a little more time in the details. He discusses reference borrowing, multiple groundings of names and various indirect forms of naming besides other issues. He even applies his theory to modal and propositional attitude contexts (Devitt 1981: 207-74). In short, Devitt has done more than anyone else in bringing the potential of the direct-causal theory to the fore.

Devitt calls baptisms and other naming situations "causal groundings". The primary form of the grounding of a proper name is perceptual contact (Devitt 1981: 133-6). For example Devitt's wife gave their cat the name "Nana". Subsequent groundings for other speakers took place when the Devitts told those people that their cat's name is "Nana", or, in the presence of Nana, they said that "this is Nana" (and presumably also pointed to the cat). By these different ways of grounding people get "grounding thoughts" about the object. The grounding thoughts are then the basis on which are built all other thoughts that involve the referent of the singular term in question. Identity beliefs help in passing on the grounding thoughts to non-demonstrative, and non-perceptual, situations. Thus if one thinks "that cat is Nana", this identity belief is likely to lead later to other beliefs and thoughts involving Nana. The person in question can also assert further statements to other people that then gain the benefit of the grounding thoughts in the absence of Nana.

Devitt explicitly adds to the direct-causal theory a component that purports to explain the continuous retention of a referential chain from an object to the subsequent uses of its name. This component is reference borrowing (Devitt 1981: 137-8). (I do not mean to imply that Kripke's picture does not have the notion of reference borrowing; of course it does, because without it the whole picture would be not a causal at all. I am only saying that Kripke did not dwell on the details of reference borrowing, so that Devitt's contribution is important.) Reference borrowing, via a proper name, from an earlier speaker is possible because the name belongs to the *designational chain* that originated in the grounding event. The notion of designational chain marks the causal links between a singular term and its referent such that the term is grounded perceptually in the referent. In reference borrowing the hearer acquires both i) the ability to refer to the object or a person that is in the starting point of the designational chain, and ii) thoughts about that object or a person. (*The ability because on Devitt's view there is a distinct ability with every proper name the speaker possesses.*) The designational chain, and networks of them, plays the primary explanatory role in Devitt's theory. Reference borrowing is, consequently, intended to supply the specific mechanism that is lacking in Kripke's rather cavalier talk of a name being connected to the original referent.

Devitt also concurs with Kripke and argues that the description theory is unable to invoke reference borrowing. - alleged appearances notwithstanding (see Strawson's discussion in *Individuals* (Strawson 1959)). Devitt's reason is that there is no guarantee that this would not lead to a circle in where speaker A has borrowed a description from speaker B, who have borrowed it from speaker C who have borrowed it from speaker A in his turn. Related to this point the plausibility of the direct-causal theory also increases when multiple groundings are taken into account (Devitt 1981:

56-7, 136). That is, typically a proper name gets linked to its referent perceptually and demonstratively, but also by someone acquiring new information about it in other occasions after the naming event by the help of different speakers. In these communicative events the possibilities of the designational references increase.

Another fruitful addition that Devitt makes is the notion of partial designation, although as Devitt mentions it was originally proposed by Hartry Field (Field 1973). There is no doubt that sometimes it is not clear whether a name refers to an entity X or to an entity Y. For example there was a time when "Madagascar" was used to refer both to the African mainland and to the island in the east coast (which is the name's current sole referent) (Evans 1977: 202). The evident explanation for this phase of mixed reference is that the causal chains were then ubiquitous and grounded in two different events (and most likely also later became multiply grounded). In this case it seems appropriate to say that "Madagascar" referred partially to the main land and partially to the island. It is easy to see that in this way the notion of partial designation can help in explaining reference change: the "first" part of the reference have later become "dropped" (forgotten, ignored) during the communicative encounters, thus leaving only the "second" part left as the sole referent of the name (Devitt 1981: 138-51).

In Kripke's picture the retention of the causal referential-cum-communicative chain intact, thus preserving the referent, is explained by the intentions to co-refer. Equivalently in Devitt's theory this is explained by the intentions to use the borrowed name by the hearer to refer to the individual the speaker refers to with the name. In light of the Madagascar example this general explanation is much too weak, however. The reference change may have taken place no matter how hard the speakers have intended to keep on co-referring. And as Devitt does realize, the intentions "seem as much in need of explanation as designation itself" (Devitt 1981: 138). But he does not try to explain them in any way. (I will take up this issue later.)

It should be realized that there seems to be no reason why the reference change through partial reference would not work for the (Fregean) description theory as well. If the causal account leads to occasional indeterminacies and relaxings of the demand of the uniqueness of the referent, the description theory must be allowed that possibility too, by parity of reasoning. This is all for the good: it happens that some descriptions are not specific enough to single out one particular referent, but few instead (but not too many, like in Kripke's Feynman example with the description "a famous physicist"). There is the "pragmatic side" of it also. Sometimes it may be preferable to leave it open, or not to disclose the referent. This usually happens when discreteness is required but it is still left to the hearer to single out the most likely referent.

To get a more adequate picture of how Devitt's theory of causal reference works, let us take a short look at how he handles the puzzles of reference, first the problem of existential statements (Devitt 1981: 186).⁷ Devitt states that both of the existentials "X does exist" and "X does not exist" are meaningful, contrary to what Russell claimed. The first existential is meaningful because there is an object or a person on which the causal links are grounded, namely the object or a person on which the designational chains originate. The second existential is meaningful because there exists a causal network which is grounded on an "imaginative act" (for instance a story).⁸ Taken at

face value these proposals are obviously acceptable; and I do not see that neither the proponent of the Fregean nor the description theory would disagree at all with them, provided the groundings are explained by descriptive fixings at least partly expressive of the sense of the referent, including the perceptual information gathered in pointing to it.

The puzzles of the trivial and informative identity statements receive the following explanation. Devitt recognizes that the direct-causal theory does need some kind of notion of sense as a mode of presentation of the referent. He proposes that that sense consists of the types of designational chains that originate from the referent of a proper name (Devitt 1981: 152-4, 236; Devitt 1989: 227-31). Accordingly the difference between the two types of identity statements is explained by the different designational chains underlying the two distinct singular terms in them respectively, both grounded in the same referent. In general the epistemic difference between the respective statements is explained by the speaker exercising two different designational abilities that are related to the two expressions "a" and "b" in the informational identity statement.

Lastly, Devitt understands the problem of the propositional attitude contexts to be that of *exportation*, i.e. under what conditions singular terms appearing in the scope of an attitude verb can be said to refer to a definite entity. Another way to express this is that the exportation goes with the validity of the existential generalization: under what conditions can we infer from the fact that X believes that *a* is F to the statement that there exists a particular individual such that it is F. Why is it that in some conditions exportation is in order, as when someone believes that the author of *Über Sinn und Bedeutung* had moustache, but it is not in order when anyone believes, say, that Santa Claus has moustache? Devitt's answer is that exportation goes through when the singular term in question, descriptions included, is non-empty and designational. That is, when the singular term has a definite referent and the causal chain that underlies it is grounded in that referent (Devitt 1981: ch.9; Devitt 1989: 229). But this is not sufficient, as Devitt himself acknowledges. A thinker or a speaker has to have the referent "in mind". This notion Devitt spells out along the by now familiar causal lines. There must exist a designational causal chain from the referent to the particular use of the singular term in question. Moreover, he tries to preserve Frege's "functionality principle" which states, according to Devitt, that the meaning of a statement or a thought depends on the meaning of its components. In relation to the propositional attitudes this means that the sense of a belief is made up by the "customary senses" of its components. So according to Devitt if you believe that Cicero was an orator, the senses are the (subsets of the) designational chains underlying the use of "Cicero" and "orator", and are grounded in them, respectively. As Devitt says "Cicero" and "orator" *specify* the modes of their reference (Devitt 1989: 238). In short, Devitt takes senses to be whole causal chains, or at least parts of them.

Notes

1. Dummett also charges Frege not keeping the different ingredients sufficiently distinct from one another (Dummett 1981: 128). The ingredients that Dummett distinguishes are in effect the same as Burge's. But in addition Dummett introduced a fourth subnotion of sense. That is sense as significance, the public meaning of a term used in communication. (I must add that it seems to be an open question if there is such a meaning, i.e. how

intersubjectively uniform it is – and/or how uniform it is required to be to serve communicative purposes.)

2. This is "persistent rigid designation" in Salmon's terminology, whereas "obstinate rigid designation" means that a name refers to the same individual in every possible world, whether the individual exists there or not (Salmon 1982: 33-4).
3. Another related distinction is that between meaning and reference fixing of proper names. According to it the meaning of a proper name can be spelled out at the level of linguistic meaning. It follows that the substance of the distinction becomes that of saying that the meaning of a proper name "NN" is that of "the bearer of 'NN'" (Recanati 1993: 155). This view is called "the metalinguistic view" by Francois Recanati and "the nominal description theory" by Kent Bach (Bach 1987). The theory of the reference of proper names is differentiated from this view simply because it would be circular to try to account for the reference of a proper name by the above metalinguistic formulation, or some of its equivalents like "who is called 'NN'" (Recanati 1993: 159-60). (See Kripke's condition (C) in section 3.5.3.) But it seems to me that the metalinguistic view is only a general theory pertaining to characterize proper names as a class of linguistic expressions *qua* referential devices. The Fregean sense theory of reference of proper names is not thereby affected, because it maintains that the referent of any particular proper name is determined, in some way or another, by the sense of that name. The linguistic meaning has nothing to do with this.
4. *Pace* those philosophers who take intuitions as sufficiently strong *evidential* base. I will not go here into the discussion of the status of intuitions as evidence for philosophical theories (but see note 29 in chapter 4). (Though there are many very problematic features of intuitions, beginning obviously from the questions *are* there intuitions, and if there are, are they sufficiently uniform at all for the purposes some philosophers insist using them?)
5. Searle seems to assume necessity in some form when he states that it is necessary that an inclusive disjunction of properties characterizes the bearer of the proper name in question; if the bearer does not satisfy at least some of those properties, there is no individual (called by that name) (Searle 1958: 172). *If* the necessity here is meant in some strong sense (maybe even implying some kind of essentialism), then I should advice that it be dropped from Searle's theory.
6. This ban on modal considerations is problematical. It does not distinguish possible worlds discourse from the common counterfactual talk in our language (for instance "I could have done that"). Therefore the accusation of the metaphoricalness of possible worlds discourse (though no doubt being to the point in some issues) does not rule out the counterfactual discourse, which suffices for making Kripke's points about rigid designators. But I let that slip pass here.
7. Devitt discusses many problematic cases of designations of proper names (Devitt 1981: 138-60). I do not go into them here for the reason that in all of them (sometimes explicitly, sometimes clearly between the lines) Devitt is forced to invoke the sense or the description theoretic factors in answering the problematic cases. Especially clear this is in the case of attributive terms. Another is the partial designation and possible reference change in the case of the Liebknichts (Devitt 1981: 140-1). It is of relevance just because the description theoretic factors can be seen to play a strong, even if somewhat tacit, part

in it. Ignorant Joe hears some of the politically acute among his friends talking about Liebknecht. Ignorant as he is, he has *keine Ahnung* that there are two relevant referents for that same name in reality as well as in those discussions, Wilhelm and Karl (father and son). When Joe then picks up the name "Liebknecht" and happens to use it, it refers partially to Wilhelm and partially to Karl. But this explanation shows that the causal chains themselves do not have the sole referential role. For consider this: if Joe were interrogated about Liebknecht, he would immediately run into troubles in being unable to give any correct answers, excluding trivial ones like "Liebknecht is a man". I want to claim that this is as firm form of evidence as we can get about the referential substance that "Liebknecht" enjoys in Joe's mouth, with the plainly negative verdict. In other words, it is the informational impact that contributes to the referential use of a proper name, not the causal links themselves.

8. This account is very similar to Donnellan's "block" explanation (Donnellan 1974).

4. CRITIQUE OF THE DIRECT-CAUSAL THEORY OF REFERENCE

The modal and the epistemological argument

Without any further ado let us take up the main arguments marshaled by the direct-causal theorists against the Fregean sense theory of reference (still with understanding that the talk of the descriptions and the description theory is also a form of the Fregean sense theory because it is held in effect by all parties that descriptions *express* the senses of proper names). Irrespective of the differences in the particular versions of the Fregean description theory the direct-causal theorists have assumed them to belong to the domain of what they have called “the traditional theories of meaning” that are thought to consist of analyticities between the predicates expressing the properties of the individuals (and the natural kinds) and proper names of those individuals (and the natural kind terms) (Putnam 1987a: 140). For example it is supposed to be analytic that, say, Aristotle was the most famous pupil of Plato. Kripke claimed in effect that “Frege-Russell view” is about the meanings in the sense of synonymies, and resulting in tautologies, when for example the meaning (i.e. what Kripke equated with Frege’s notion of sense) of “Aristotle” is expressed by “the man who taught Alexander the Great” (Kripke 1980: 30, 32, 58). Sometimes he also seems to presuppose that the Frege-Russell view requires only one particular description expressing the meaning/sense of a particular proper name, and saying that Searle’s cluster view is the *locus classicus* of the more adequate *description* theory (Kripke 1980: 31). And related to the semantic point of view, Kripke says explicitly that Searle’s view “...may seem, as an *analysis of ordinary language* quite a bit more plausible than that of Frege and Russell.” (Kripke 1980: 31; my emphasis) But why then, in the Preface in (Kripke 1980), does he concentrate on Russell’s view as a single-description per name view? That is strained practice, even if not incoherent. I have no firm answer to that question; it might be that Kripke was only careless in his 1980 formulation for the reason that meanwhile, since 1972, it became a common view not to draw any distinctions between the various versions of the description theory and Frege’s original account; hence the general label “Frege-Russell theory”.

However I suggest resolving this discrepancy by simply assuming that there just is no relevant difference between what Kripke calls “Frege-Russell theory” and the cluster version when it comes to the theory of the *meaning* as well as theory of the *reference* of proper names. At least Russell never held the view that there is only one description for a particular ordinary proper name; and Frege explicitly denied that view in his famous Aristotle note where he says that to a particular proper name relates various senses. Another important point is that even if Searle held his cluster theory responsible for involving analyticities and synonymies, but because neither Frege nor Russell did so, the latter views remain sound and viable even if Searle’s cluster version succumbs.

Kripke introduces the distinction between giving the meaning of a proper name and fixing the referent of that name (though he mentions Paul Ziff’s account of proper names in this connection). He points out correctly that if a name could not be analyzed by a description, because they are not synonymous, but only “determined” by a description, proper names are in general only “materially equivalent” to the descriptions (Kripke 1980: 33). But Kripke leaves it open whether or not material equivalence amounts to the description theory being an empirical theory about the

functioning of proper names. In pages 53-4 Kripke states that the contrast is between the theory of meaning of names and “merely” the theory of the reference of proper names. This would seem to indicate that Kripke understands the theory of reference to be an empirical theory, or an account at least (“picture”), not a philosophical theory - which he denounces and proposes instead his causal picture of the reference of proper names. But, again, he does not say much such that would enable one to state firmly that a picture of reference is an empirical theory by his lights. However I propose to take it that way, for what Kripke says about his picture is what anyone explicit in his intention would say if he were to offer an empirical theory of reference. Consequently I also take it that the Fregean theory of sense and reference – because it is *not* a philosophical theory, i.e. it does *not* attempt to *analyze* proper names by the definite descriptions or give necessary and sufficient conditions for reference (except trivially that if there is an individuating and identifying sense, there is reference) – competes with the direct-causal theory in that both are offered as empirical accounts of the reference of proper names.

As we will now see the modal and the epistemological arguments can be ruled out as attacks on the Fregean straw man. Although it has become almost unanimously accepted that according to the Fregean description theory proper names and the definite descriptions are synonymous with one another and that the identity statements in which they appear are analytic statements (and definitionally equivalent), such demands were not part of Frege’s theory. It is true that Frege wrote about defining non-primitive terms, especially mathematical ones, and the analyticities between them through definitions. But then his focus was on mathematical objects and on the concept of number. His logic and *Begriffsschrift* were to provide the base for mathematics and also to be the format for the ideal scientific language, in which no expressions occur which have no *Bedeutungen* (although they have senses). Frege himself pointed out not a few times that the expressions of natural language are ill suited for such treatment, for, among other things, natural language contains expressions which do not have *Bedeutungen*. And when it comes to the description theorists recall Searle’s correct observation that we do not define proper names in natural language (Searle 1958: 166). In addition we find from the neo-Fregean camp David Bell’s clear maintenance that the sense of a proper name is “the condition which something must meet if it is to be the referent of ‘a’, i.e. if it is to be *a*...It does not, however, follow from this theory...that the proper name in question is synonymous with any expression which applies to the appropriate condition.” (Bell 1984: 184-5) (see also (Dummett 1981: 341-2)). In light of these remarks the direct-causal theorists, in particular, could not complain about the lack of synonymies between proper names and the definite descriptions because their main focus is on expressions of natural language, not on ideal scientific constructions. (And it would be open to debate whether or not synonymy should be demanded of the relevant expressions in an ideal language of science either.)

Moreover the demands for synonymies and analyticities are such strong semantic demands that clearly the burden of establishing them as reasonable requirements on the sense theory of reference is on the direct-causal theorists. But all that they have done is to have claimed and assumed that the Fregean description theory is committed to such strong semantical equivalences between proper names and the descriptions. For example Putnam did not give any textual evidence; he only says that Rudolf Carnap formalized these traditional theories, and that the theories Jerrold Katz and co-

workers have produced share the defects of the traditional theories (Putnam 1987a: 139). He also includes Plato and Frege in that tradition (Putnam 1987a: 150). But if there were many such “traditional theories” one would have thought that textual evidence would have been easy to amass. Concerning this lack of textual evidence Salmon writes that Linsky, who denies that Frege and Russell had any synonymy demands (see note 1) and criticizes Kripke for not giving any textual evidence for his contrary claim, fails himself to give any supporting textual evidence for his claim (Salmon 1979: 445, n.7). But I claim that strictly this charge from the lack of textual evidence is out of place here. I suggest that because Frege and Russell did not say that the ordinary proper names are synonymous to the definite descriptions that itself *is* evidence that they did not demand such synonymy relations. That they did not say anything like that is most likely explained by the fact that back in those days no such semantic approaches were practiced in the philosophy of language and thought; so Frege and Russell *needed* not to deny explicitly any such demands. (That would have been curious indeed.)

Russell’s theory did not demand or imply any synonymy or analyticity demands either. He held that the ordinary proper names are abbreviated (“truncated”, “telescoped”) descriptions, but that does not imply - let alone entail - synonymies or analyticities between the descriptions and the ordinary proper names. As we saw Russell said that propositions like “Scott is the author of *Waverley*” are not tautological, hence presumably not analytical and not expressing synonymies.

“If you understand English language, you would understand the meaning of the phrase ‘The author of *Waverley*’ if you had never heard it before, whereas you would not understand the meaning of ‘Scott’ if you have never heard the word before because to know the meaning of a name is to know who it is applied to.” (Russell 1956: 244)

If Russell held that the sentence is still somehow analytical and/or expresses a synonymy relation between “Scott” and “the author of *Waverley*”, most likely he would have said so, and would have elaborated this in relation to his theory of the definite descriptions. But that seems not to have been the case.¹

It is true that in “On Denoting” Russell wrote that “...when there is anything with which we do not have immediate acquaintance, but only definition by denoting phrases...” (Russell 1956: 55). But it is clear that for Russell “definition” did not mean synonymous paraphrases, or any such constructions. Rather “definitions” is to be understood as a characterization given by the definite description(s) used, no matter how many or few of them are needed for a unique characterization of the object of the denotation to obtain. Uniqueness is the point, not semantical relations like synonymies or analyticities. Neither did Russell say anything about such strict semantical relations as holding between the ordinary proper names and denoting phrases in “On Denoting” or elsewhere. Moreover the general perspective in the place from where the above quotation is drawn is epistemic: the issue is knowledge by acquaintance – knowledge by description. So not even here it is required by Russell’s theory that the “definition” by the descriptive phrases is once and for all locked in with the proper name in question; only that what is known about the “terms”, i.e. the referents, with the help of the phrases could change (because as far as the epistemic issue goes, we could be in error with respect to the denoting phrases).

It is also true that Russell says, when talking about “non-entities” like round squares, Apollo and Hamlet, that “A proposition about Apollo means what we get by substituting what the classical dictionary tells us is meant by Apollo, say the ‘sun-god’.” (Russell 1956: 54) But, again, “means” coupled with the casual “say the sun-god” makes it evident that meaning is not meant in the strict sense of synonymy. “Means” just means that the denoting phrases (from dictionaries and encyclopedias, and other such information sources, like plainly hearing someone telling something related) give us information about the denoted objects, descriptions with which we come to know what the objects are (as they are characterized by that information). In particular Russell does *not* say that the ordinary proper names are defined to be synonymous with the descriptions found in the dictionaries and such sources. (Though I have come across with such extreme interpretations – deliberate misreadings?) It is true that Russell speaks of defining propositions like “the author of *Waverley* was Scotch” by clauses amounting to “at least and most one person wrote *Waverley*, and that person was Scotch” (Russell 1993: 53). But again this is just that: defining *propositions*, not proper names with one description per name. A similar “definition” can be given by any property of the person in question. (If one likes one could take all true propositions of that person to constitute, then, the definition of the name, but that would amount to the Leibnizian perfect individual concept.)

The view that Russell is not restricted in his theory to strict definitions receives further support – maybe even conclusive evidence – in a passage where he discusses the distinctions between the ordinary proper names and the descriptions which is motivated by his metaphysics and epistemology of logical atomism (Russell 1956: 200-1). The logically proper names are words for particulars with which we are only acquainted, and the *ordinary* proper names are abbreviations for the descriptions because we use those names for objects we are not acquainted with. At least not always: “Socrates” could not be but an abbreviation for *us* since we could not be acquainted with him for the obvious reason. Our thoughts about Socrates can be rendered by any common phrase like “The Master of Plato” or “the philosopher who drank hemlock”. That Russell is neither giving nor demanding any strict definitions but only “renderings”, i.e. loose relatedness of the descriptions with the names, should make it plain that Russell’s theory of descriptions has nothing semantically essential to do with synonymies and such equivalent notions.

Besides, when Kripke argues for the difference between proper names and the descriptions he compares statements like “Aristotle is Aristotle” to statements like “the most famous pupil of Plato is Aristotle” and says that the former is necessarily true but the latter is not. Apart from the general modal emphasis in Kripke’s way of arguing, how does this differ from what Russell argues for example that “the author of *Waverley*” is different from “Scott” because if it were not, “Scott is the author of *Waverley*” would express the same proposition as “Scott is Scott” – which it obviously does not (Russell 1956: 252-3). So it can be argued that Kripke’s view is not so anti-descriptivist at all in this respect. (Except with respect to Searle’s theory, but even that is not quite clear because what Searle really means by “necessary conjunction of propositions” is unclear. Though Searle writes in another place in a way that seems to commit him to analyticities: “it is a necessary condition for an object’s being Aristotle that it satisfy at least some of these descriptions. This is another way of saying that the *disjunction* of these descriptions is *analytically tied* to the name ‘Aristotle’.” (Searle

1967: 490; emphasis in the original) If the difference between proper names and the descriptions is there even in Russell's theory, one is left to ask what is the issue about the rigidity of proper names, as something that the description theory of reference (or the Fregean theory for that matter) could not contain in its theoretical tool kit. However I will postpone this issue for a while and discuss it more thoroughly in section 4.5 and in chapter 5.

At this point I conclude that with respect to the Fregean sense theory of reference there are no grounds for the modal argument: whatever semantical relations there are between proper names and the descriptions associated to them, they are not relations of synonymy, neither of analyticity nor of logical truth. The epistemological argument could be expected to be without such grounds also because it is a close kin to the modal one. And indeed that is so, and for similar reasons. It is quite correct as Salmon states that the sentence "Shakespeare wrote *Hamlet*" is not knowable *a priori* by the concepts involved in it, for by themselves only they do not make it the case that only Shakespeare wrote the plays and did what he did (in all possible worlds actual world obviously included). But then this has been claimed by neither the sense nor the description theorists. One does not find evidence for this strong view in the writings of those theorists. On the contrary Russell held, just like Frege did, that it is *precisely* because we can *not* attain a complete knowledge of the referents, such that would enable us to have Leibnizian complete individual concepts of them and turn the above type of identity sentences into logical truths, that the ordinary proper names are truncated descriptions and incomplete symbols. (See (Frege 1997a: 153) that was mentioned in section 2.2.) Moreover the epistemological argument seems confused. Note that its *conclusion* is just that what forms the *starting* point of Frege's puzzle about trivial and informational identity statements. Because "a is b" is informative, i.e. it is not known *a priori* by the concepts involved in the names and the descriptions used, there is something in it that is in need of explanation. The identity puzzle itself shows that the Fregean theory of sense and reference has nothing to do with any *a priori* demands between proper names and the descriptions.²

In analogy to the familiar distinction between the semantic reference and the speaker reference we may adopt the distinction between semantic sense and speaker sense. The direct-causal theorists have uniformly assumed in their critiques of the Fregean theory the former perspective, that of the semantic senses, the senses associated uniformly to a proper name in the linguistic community as a whole. But that is a mistake for it is the perspective from the speaker senses that is relevant: remember again Frege's Aristotle note where it is allowed that different persons have different senses associated to the same proper name. Within this perspective the senses *can* be taken as synonymous to descriptions (expressing the senses). But then it should be realized that this move is also inconsequential because it is temporally relativized to the particular occasions of the uses of proper names by individual speakers. For in another occasion the same speaker may invoke other senses associated to the name. Of course this temporalization means in effect that proper names are not synonymous uniformly with the same senses from an occasion to another. Idealizingly, though, one could argue that the set of all the senses the speaker invokes (consciously as well as tacitly) in all occasions he uses the name, the total "sense set", is synonymous to the proper name for that speaker. But the sense theorists need not assume that.

Argumentatively the issue can be left here. But it seems to me that because the gap between the demands imposed by the direct-causal theorists on the sense theory and the actual substance of that theory is so wide, an explanation of the gap is in order. I will make only a suggestion here and not pursue it any further. (Though of course I think that it would make an interesting case for both a substantial and a sociopsychological study of the history of the analytic philosophy of language (especially in US) during the latter part of 20th century.) It seems to me that back in the 1960s there was very much in force a presupposition, common and therefore mostly tacitly accepted, about the general methodology in philosophy and in semantical studies particularly. According to that presupposition meaning relations were taken to amount only to synonymities, analyticities, meaning postulates, definitions, explications (in the Carnapian style) and other such strong semantic notions. The general methodological atmosphere in the analytic philosophy in those days, after the "linguistic turn" effected by the logical empiricists, presumably was such that philosophers were trying to offer conceptual analyses of many interesting notions and providing definitions of the central terms in the manner that would preserve the meaning of the *definiendum* by the *definiens*, and *analysandum* by the *analysans*, i.e. to provide synonymy-preserving analyses. The analyses were to give the conditions of application for being (the extension of) the concept in question. The conditions given by the *analysans* and *definiens* were supposed to give *the* meaning both of the concept and the term that linguistically expresses the concept. The prevalence of this methodology in those years is representatively evident in Arthur Pap's book with a telling name *Semantics and Necessary Truth. An Inquiry into the Foundations of Analytic Philosophy* (Pap 1958). Pap writes about the relationships between proper names and the descriptions and maintains that proper names are defined by the descriptions, and that for proper names one should find such description which is synonymous with it

"...any proper name is synonymous with a definite description...("individual concepts", in the present terminology of Carnap): for if it makes sense to speak, say, of the concept of Caesar, then it ought to make sense to speak of *defining* the proper name "Caesar", but this can only mean to produce a definite description with which the name is synonymous." (Pap 1958: 18)

But still as recently as 2002 and 2005 the direct-causal theorists, or the Millian theorists, work under the misconception of synonymity. An instance of this is Scott Soames in his two recent books on the issues of semantic values and references of proper names versus the definite descriptions (Soames 2002, 2005). Soames explicitly assumes that it is the core claim of the description theoretic version of the sense theory of reference that proper names are intersubjectively invariantly synonymous with the definite descriptions related to them semantically.

I argued above that there is no textual evidence for the synonymy interpretation. In case the reader thinks that I need more substantial or systematic argument to show my case against synonymies, I offer the following. Frege held that the referents of the number expressions – "1", "2", "3" and such – are the values of the different functions having them as the results of the applications of the functions. For example "2+2=4" and "2²=4" involve two different functions (2+2= and 2²=) but one and the same number as the result. In other words, for any number being the referent, there is many

(even possibly infinitely many) ways it can be “grasped” via the corresponding functions (analogous to the senses). So it would be very odd to claim that any linguistic expression of one of these functions gives the synonymous expression (description) for that number. That “2+2” is, perhaps, the most often used function for representing 4 is only because it is just that: most often used expression for 4. There is nothing more to it semantically, so “2+2” is not synonymous for “4”. In the same way a single description, expressing the sense, can be often used to characterize a certain individual, but it does not follow that that description is the synonymy expression for the proper name of that individual. Neither needs the sense that description expresses be intersubjectively the sole sense related to the individual as a mode of its presentation. Therefore I conclude that the synonymy demand for proper names has always been a (harmful) mistake made by the direct-causal (and Millian) theorists.

That Kripke endorsed the synonymy approach, and assumed that it is the core of the Fregean sense theory, or at least should, becomes clear when he differentiates between the theory of the meaning of proper names and the theory of the reference of proper names (Kripke 1972: 258-9). He also implies that definitions and analyses of proper names is the business of the theory of meaning, and attributes this aim also to the Searlean cluster version of the description theory (Kripke 1972: 276-8). Of Frege and Russell he claims that they “certainly seem to have the full blown theory according to which a proper name is not a rigid designator and is synonymous with the description which replaces it.”³ Hence the motivation behind the modal and epistemological arguments against the sense theory of reference become comprehensible. From this conception it was also a short step to considering the senses of proper names as their meaning conditions. Therefore the Fregean senses which have not just semantical tasks to perform as intensional entities, but also epistemological and cognitive tasks, were immediately put within the confines of that restricted conception of semantics. In other words the direct-causal theorists ignored the rich cognitive and epistemic roles that the senses (and the thoughts) became to assume in Frege’s overall view. Nowadays, when the notion of meaning, or content, has become to be taken as cognitively and epistemically respectable notion, we have got used to the idea that meanings, and semantic matters in general, can not be restricted in the manner that was the dominating approach still in the 1960s. It was also part of the methodology of the conceptual analyses to give such conditions for concepts that could be immune from any counterexamples. That is, the conditions for being X should apply not just to actual cases but to any possible cases also. This largely explains the use of the possible-world counterexamples since Kripke.

Kripke claimed that if the sense theory is taken as a theory of the reference of proper names, but not as a theory of their (synonymous) meaning, it loses much of its appeal, especially in relation to the negative existential statements like “Moses does not exist”. But how so? The very supposition that the direct-causal theorists have of the sense theoretic solution to the negative existentials is something that the sense theorists need not accept (and have not done so, excepting Searle), let alone that it would imply that solution. Rather it is a straw man constructed by the former theorists. The negative existentials are only one problem case, and the sense theory can provide an adequate solution to them, a solution that *allows* that it can turn out that an individual NN has not any (nontrivial) properties it has been thought to have. This only requires that some new and true information be found out about that individual. But still, according to the Fregean sense theory proper names have meanings through their senses, only

those meanings do not provide synonymies between the descriptions that express the senses and proper names of which they are the senses. The most general form of this argument is that the sense theory is independent from the notion of synonymy: one can be a "synonymy nihilist" and still a whole hearted Fregean with respect to proper names. In other words there is no conceptual connection from the sense theory to the synonymy properties of proper names - unlike it seems to be the case with the conception of semantics the direct-causal theorists have adopted.

We will see in section 4.4 that the modal, epistemological as well as the semantical argument are all based on a presupposition about the individuation of the referents of proper names which the direct-causal theorists always make but which in fact destroys the whole notion of proper names as rigid designators independent of the Fregean senses as the factors which individuate their referents. This is not to deny that proper names and the definite descriptions sometimes behave differently in the intensional contexts. But this will disarm at large the direct-causal theorists and will no more enable them to provide any modal arguments against the sense theory of reference via the notion of the rigid designation. (I will also explain the differential behavior of proper names and the definite descriptions in chapters 5 and 6.)

As I already pointed out Kripke belittles the sense theory when it is taken "only" to be a theory of the reference of proper names and not of their meanings. His main reason is that as a theory of reference the sense theory loses much of its appeal in dealing with the negative existential statements. He also seems to maintain that as a theory of reference the sense theory is equivalent to the account in which descriptions *fix* the referents of proper names. According to Kripke the negative existential statements ("X does not exist") can not be analyzed to mean that if the descriptions are not true of anyone the proper name in question has no referent. But this charge is hard to take seriously, and it seems to me to be incorrect. The reason is that surely the sense theory, as a theory of reference, is not to be diminished to offering mere fixing role to the senses. The sense theory of reference is an *empirical* theory, not an old-fashioned semantical theory offering synonymies and other such stringent semantic relations. Therefore ridding it of the philosophical interpretation (the methodology of meaning analyses) leaves it every opportunity to explain the mechanisms and factors effective when proper names are used referentially. Moreover the sense theory of reference does explain the negative existentials. The statement "X does not exist" simply states that "X" does not refer to any actual individual. But "X" has sense, and that accounts for the fact that we did believe that it referred until the referential emptiness of "X" was discovered. Consequently Kripke's charge from the weakness of the sense theory as a mere theory of reference should not be taken at all seriously (neither should any Fregean theorist with an empirical bent take it so).

To continue a little more on the issue of analyzing the meaning versus referring in relation to the account of fixing the referent, the latter option fulfills the functions of the three subsenses described earlier. For in fixing the referent by a description (or descriptions) the speaker associates the sense(s) the description(s) express(es) with the proper name, this being the *conceptual* representation of the referent. The speaker also *determines* the referent by the description(s) just because he uses it (them) to fix the referent. And finally the sense the speaker expresses with the fixing description(s) is the *information content* which he contributes to the sentences in which the proper name occurs, also in the intensional contexts.⁵ No intersubjective sameness of content

is required, *pace* Burge and Salmon. Therefore what is left for the sense theorist to be concerned about the direct-causal theorists' arguments is the possibility that the description (or set of descriptions) turns out to be false, but that the proper name still allegedly refers. I will also deal with this residual argument later. But here we can conclude that the charge from the distinction between the theory of meaning and the theory of reference fails with respect to the Fregean theory of sense and reference.

So the modal and the epistemological arguments fail as arguments against the Fregean and Russellian theories because they are not committed to any stringent semantic demands like synonymy and analytic meaning equivalences or *a priori* knowledge between proper names and the definite descriptions. But there is another way to read the arguments, and it seems to be the real intention in Kripke's discussion. Or at least the arguments could be taken this other way. The stringent semantic desiderata can be dropped but the modal argument can still be used in its essentials. We made acquaintance with that way when the notion of rigid designation was explicated. To rehearse it, take any individuating sense expressed by a definite description, for instance "the most famous pupil of Plato". This refers to Aristotle. But in some counterfactual situation (a possible world) the referent of that description might *not* be the actual Aristotle but some other individual altogether. Still, the proper name "Aristotle" refers to the *actual* Aristotle in those circumstances. In other words proper names and the descriptions behave differently in modal contexts. But this argument can be rebutted – and it was rebutted by Alvin Plantinga only a few years after the publication of Kripke's 1972 published lectures (Plantinga 1978). The rebuttal goes by the name "actualized descriptions" (or some such equivalent). According to this view the definite descriptions related to proper names become rigidified when an "actually" operator is attached to them. However I shall postpone that issue until the next chapter. The reason is that there I will offer a new look at the actualized-rigidified theory as well as on Kripke's modal argument.

The semantical argument

The status of the argument

The semantic argument receives a long treatment in the following sections. Therefore it is in order to present the general course of my argumentation. First I argue, using the well-known case of Madagascar that the causal links by themselves are not sufficiently powerful to guarantee reference such that the original referent of the name would be guaranteed to remain retained through the subsequent exchanges. After that I argue that the intentions to co-refer that the direct-causal theorists have realized they need to postulate do not engender sufficient referential force either. Furthermore the usual types of examples used by the direct-causal theorists – the plausible Feynman case and the counterfactual Gödel case – are defective in that they seem to invoke tacitly informational elements in a way that the correct diagnosis drawn from the examples is not what the direct-causal theorists claim it to be. This does not leave those theorists totally disarmed, however. They can still claim that the sense and description theories fail because there are cases when the descriptions (expressing the senses) are false of the referent of the proper name in question but despite that being so the name still refers successfully. Indeed when the misconception about the senses are put aside and Kripke's argument is interpreted this way it presents the core claim of the direct-causal

theory: proper names do refer “on their own” if the communicative causal links are there. This can be put in general terms along the lines that Robert Stalnaker has presented the theoretical situation: the direct-causal theory involves a semantic claim and a metasemantic claim. The semantic claim is that the sole *semantic value* of a referring expression is its referent; the metasemantic claim, or the metatheoretical claim, is that the referent being the semantic value is explained by it standing at the starting point of the appropriate causal links to the subsequent uses of its name. This leads me to scrutinize the much-used intuitive pull behind such claims. Consequently I will argue that the intuitions behind the direct-causal claim are misconceived: they are not semantical, *a fortiori* not substantially referential, but relate to the *preconditions* of all referential language use. Dialectically the situation is as follows. We begin with the current consensus according to which the Fregean and the description theories are false, but the direct-causal theory is correct (at least more clearly so than the former theories). Then along the course of my argumentation I strip the latter theory of its causal clothes, the result being almost the total reverse of the current consensus: the direct-causal theory is left with only the defensive move of charging the Fregean and the description theory that they could not handle the case where the descriptions are false but the proper name in question does refer. But even that straw breaks in the end, as we will see.

The semantical argument does not presuppose any synonymy or analytic demands between proper names and the definite descriptions. Instead it is directed against the idea that the descriptions give the necessary and sufficient conditions for the referent of a proper name. Searle’s cluster version has been the main target but the argument generalizes obviously to the Fregean sense theory. To rehearse the argument in its general form a proper name refers to that individual from whom the uses of that name originated and are derived via the causal communicative chains. The name does not refer to any occasional individual to whom the description applies.

As I already pointed out Salmon emphasized that this is a clear answer to a clear semantic question: to what we would refer were the envisaged situation to occur? In stating this question Salmon relies on a notion he calls “reference *simpliciter*”, which, as judged from his example and wording, presumably relates only to proper names but not to the descriptions. It seems to me that in this respect it is similar, if not even the same, as Recanati’s notion of REF that we already encountered and by which the latter tries to account for the referential directness of proper names (again in contradistinction to the descriptions, other than the rigid ones). So I suppose that Salmon’s argument is intended as some kind of *conceptual* argument about the intrinsic nature of the concept of reference. (However, and contrary to what I just surmised, about Recanati’s intentions with REF I am not sure.)

I am not happy with Salmon’s interpretation of the semantical argument, neither with the notion of reference *simpliciter*. It seems to me that the intuitive appeal that Salmon claims on behalf of the exclusive referential success of the name “Thales” just is not there. Or if it is there, it is relative to a given speaker. For example my intuitive grasp of the Thales case gives the contrary verdict, for it seems to me that if “Thales” refers to anything in the situation envisaged, it *does* refer to the water metaphysicist, not to the Thales from whom the name originated. For me it is obvious that the description that expresses what the metaphysician believed (and what I and, by large, we believe of that chap) outweighs the origin of the name, with respect to referential intuitions. As

for me I believe that Thales lived in the 6th century before Christ; was from Miletos; traveled in Egypt; predicted an eclipse of the Sun; held, indeed, that water is the *Urstoff*; that Diogenes Laertius and Aristotle wrote about him; that he is said to have fallen into a well while gazing stars. To this adds the fact, which Salmon does not rule out but to which he in fact seems implicitly committed in his endorsing to achieve the alleged contrast effect of the example that "Thales" has been used with discussions and informational exchanges the topics of which have been mainly, I suppose, metaphysics and Greek philosophy. In other words, "Thales" has become so tied to this informational, cognitive and epistemic background that the supposed original Thales has dropped out of the picture very early on. (That is, were there such an individual the example envisions to make its point about.)

We need not be restricted to my contrary intuitions to show that the Salmonian reference *simpliciter* is not a uniform notion. Let us take a well-known case in addition. Linsky discusses "the Homeric question", the problem whether one person wrote the Homeric Poems or are they perhaps a compilation from different writers (Linsky 1977: 96-7). Now *if* the Poems were written by a single person, we could construct a case in which the name "Homer" derives from an ancient Greek with that name but who did not write anything. Even so the Homeric question, because it has been pondered over so long, have created a strong "epistemic surrounding" to the effect that when we use "Homer" we are intending to talk about the writer of the Poems, whoever he or she was, and whatever his or her name was. It seems to me that, unless one is thoroughly immersed in the direct-causal picture, one realizes even after a quick reflection on this example that "Homer" is used in the way that its referent is the individual who wrote the Poems, irrespective whether or not there was some other individual named "Homer".

A contrary intuition is recorded in (Evans 1982: 395-6). (But see also (Evans 1982: 378, n. 6) on "Jack the Ripper" and "Deutero-Isaiah", as well as Evans' famous "Julius", the whoever invented the zip, all of which seem curiously incompatible with Evans' own intuition.) In explaining his intuition, which he takes to be the uniformly accepted one, Evans denies that we are using "Homer" as the name of whoever was the author of the *Iliad* and the *Odyssey*. But because we *are* so using it, or some of us (at least some of the historians) really are so using it, I maintain that Evans' contrary intuition, supported by the alleged uniformity, so to say counts itself out. Evans also defends his view that "Homer" is not our name for whoever was the author by maintaining that upon the possible discovery of that individual, i.e. that individual having been so-and-so and named "Homer" (or rather the ancient equivalent), we will be in a position to justify that we have been using "Homer" as the name of that very individual and not the writer. But that is not denied by anyone concerned - or the very least it should not be denied. Of course if such a discovery will be made, we may disambiguate "Homer" to refer to the discovered individual also, or we may begin to use "Homer" exclusively as the name for that individual. These are possible moves *just because* we currently seem to lack any definite and specific information about the actual author, apart from that he or she wrote the Homeric poems, and therefore we use "Homer" as the name of whoever was the author. In a way "Homer" has become a placeholder for the person whoever wrote the Poems (if such there is). It functions as a placeholder because it depends on the information, expressed by the descriptions we associate with it.

Let us take one more example, this time from the nearer history. Petrarca first met Laura morning 6.4.1327 at the church of St. Clara in Avignon, and fell in romantic love with her at once. After that Petrarca wrote the famous sonets to Laura (*Rerum vulgarium fragmenta*). But as a matter of fact we do not know whether Laura's name was "Laura" or not. Let us suppose that there was some other woman named "Laura" from whom the later uses of "Laura" derive causally. Still it seems that this Laura has not been referred to when the scholars and literature people have talked about Laura, Petrarca's love. All that has been said of Laura, mainly on the basis of the contents of Petrarca's sonets, is about that woman Petrarca first saw in Avignon. Here our intuitions overrule any original causal contact with some now unknown "Laura" in favour of the descriptive information individuating Petrarca's "Laura". (The "causal Laura" here plays the role of the "Thales" in the original example, and "Petrarca's Laura" the role of the water metaphysician to whom the descriptions apply.)

The general conclusion I draw is, then, that it is very problematical to base one's argument on an allegedly uniform notion of reference *simpliciter*, for there clearly is no uniformity with respect to the referential intuitions.⁶ The issues has to be decided on empirical grounds, not on conceptual grounds backed up by one's intuitions concerning examples. That the intuitions are much too varying and weak can be shown also by the following argument. Let us suppose that we have ignored any contrary intuitions to the direct-causal view. When we use a proper name to refer to an individual that to all intents and purposes we take to exist (or to have existed), we can detect the same referential intuitions *simpliciter* as in the Thales example. But what are the direct-causal theorists to say when it eventually turns out that the name used is nevertheless an empty one, i.e. that there is not, and has never been, any individual to whom we used the name (purportedly) to refer? And this is so especially when there is no causal link from the purported individual to the uses of its name, simply because that individual does not exist. The referential intuition is there, and it is as strong as any such Salmonian one, but the very theory that intuition is used to support must be denying its force. The obvious irony here is that the sense theory explains the referential practice and the intuition: we believed that the individual exists, or existed, because we associated senses, expressed by descriptions, with that name.

I think that we can conclude that proper names do not have any distinctive function of being referential devices *simpliciter* such that the descriptions do not have or could not become to acquire, depending on the informational situations in which they are used. I maintain that the descriptions about the water metaphysician are as strong referential devices, even intuitively, as proper names. I will argue this more thoroughly in section 4.2.5. As I already stated, another important conclusion that can be drawn from this moral is that the semantical argument is not purely conceptual argument, but an empirical one.⁷ This can be seen more clearly if we consider the point of the moral I drew according to which proper names and the descriptions do not differ from each other with respect to their indications of the referential "modes" they are capable of assuming. Accordingly to claim referential advantage for the proper name in the Thales example must, then, involve something else than the mere conceptual notion of reference *simpliciter*. And it is easy to see that this is the case. The additional factor by which the name gets to the subsequent uses is evidently causal in the example, so the notion of reference *simpliciter* is not so simple a thing at all; it is a causal notion. Therefore it is very much open to debate whether the Thales example (and others like it) can be used as part of a conceptually disguised argument against the sense theory of

reference. The reason is that in the form the example is given and interpreted as exhibiting a conceptual element in the notion of reference it *assumes* the correctness of the direct-causal account. However it should be noted that the argument is not outright question begging because the causal involvement is not spelled out in any detail, not in the example neither in the ensuing argument. Therefore we need to provide further arguments on behalf of the Fregean sense theorist (and the description theorist) and against the direct-causal picture. But to take a dialectical stock: the modal and epistemological arguments are inadequate because they have presuppositions that neither the Fregean sense theory nor the Russellian description theory are committed to, and the semantical argument will fail also if the causal determination of the referents of proper names will turn out to be too weak.

Arguments from ignorance and error

Kripke formulated the Searlean cluster version of the description theory in the way that it concerns the perspective of an individual speaker. As we remember the first argument tries to show that the senses expressed by the descriptions are not necessary for successful reference; this is the Feynman example. If someone knows only that a person called "Feynman" is a famous physicist, then according to the sense theory he is not referring uniquely to Richard Feynman because there are many famous physicists (and maybe more than one named "Feynman"). But Kripke's claim is plainly that in spite of this the speaker *is* referring to Richard Feynman, the Nobel laureate for inventing the renormalization method for QED (together with Tomonaga and Schwinger). In this case the speaker reference is on a par with the semantic reference, but in a negative way: according to Kripke's argument the speaker could not avoid being referring to any other person than Richard Feynman because that is the way that name refers in the linguistic community to what the speaker belongs. Semantic reference is assumed to overrule the speaker reference. To restate Kripke's explanation of this the causal communicative links have established this semantic practice. However it is left totally open how this could come about, i.e. what is supposed to prevent "cuts" or "mixings" taking place during the evolving courses of the causal links. (Kripke realizes this *lacuna* but nevertheless does not give any solution to it.)

The Gödel example is designed to show that the senses are not even sufficient for successful reference. Here the focus is clearly on the semantic reference as in the previous argument, for Kripke claims that *we* do refer to Kurt Gödel by the name "Gödel" and not to Schmidt, even if it were the case that the latter solely discovered the incompleteness of arithmetics, and we associate the sense expressed by the description "the discoverer of the incompleteness of arithmetics" to the name "Gödel". So again, if one intends to refer to the discoverer of the incompleteness of arithmetics by using the name "Gödel", he fails because he refers instead to Kurt Gödel. The explanation is again the weaving of the net of the causal communicative links in which "Gödel" has been used since the baptism of Kurt Gödel back in Brno. So it certainly seems that Kripke has refuted the Fregean sense theory from the semantic reference perspective, if that theory is taken to concern speaker references only.

As already pointed out Devitt ignores counterexamples that are given in terms of possible worlds. Instead he gives three types of empirical arguments. The first states that there are cases in which we do, or could, not associate an appropriate description

to a proper name, but the name does still refer to the very bearer of that name. The second argument maintains that the description identifies a wrong object as the referent of the name in question. The third argument faults the Fregean description theory with cases in which a description identifies no object but the name does refer successfully. As I already pointed out the first argument is very similar to Kripke's argument that the descriptions are not necessary to successful reference, and the second argument is similar to Kripke's argument from the insufficiency of the descriptions to successful reference. Consequently I will pack them together when I will criticize them later. Devitt's third argument deserves an independent handling, also later in this chapter. But this should be done with effect, so it is in order first to critically discuss the mechanisms of the causal reference.

Intentions to co-refer, reference borrowing by abilities to refer and the poverty of the causal chains

Is the direct-causal theory's explanation of the reference of proper names sustainable? To be able to answer this question one must study more closely the communicative links on which the theory relies, for the crucial explanatory factor of the theory is the mechanism by which the referent is retained in the linguistic practices. In Kripke's picture this mechanism is realized by the intention of the hearer to keep on referring to the same individual as does the speaker from whom he acquired the name. In Devitt's theory the mechanism is the reference borrowing. It should be realized that the intentions to co-refer and the reference borrowings are not equivalent notions, for one may borrow reference unattended without specifically intending to co-refer, and one may be intending to co-refer as sincerely as one could but still not be borrowing the reference – as is the case with the first misreference in the Madagascar example.

So the explanatory success of the direct-causal theory depends on the claim that the communicatively built links and chains are uniformly able to preserve the individual originally named. Certain uses of "Nana" refer to the Devitts' cat Nana because the uses of that name derive from the naming "ceremony". In Devitt's version this is so because the other people acquire *abilities* to refer to Nana upon hearing (or reading) about Nana. When they subsequently use "Nana" they intend to co-refer to whom or to what the person, from whom they picked the name, used the name to refer, thus borrowing reference. This allows that the descriptions used in the original naming, or the descriptions taken into use later, may be false of Nana without that severing the referential connection.

I intend to argue now that the intentions to co-refer are not capable of explaining referent retention. Let us begin with what Alan Berger says

"How, then, is a description used to transmit the reference of 'Aristotle'? The description is used to indicate something the speaker or the linguistic community, by and large, *believes* true of the referent of the term. By the use of a description, then, a speaker helps to indicate which historical chain (of intentions to corefer) he is on when he uses a given occurrence of the term 'Aristotle'. The historical chain determines the referent of the name 'Aristotle', but the use of a definite description helps indicate who the speaker or the linguistic community believes is the referent of the term." (Berger 2002: 15; an earlier statement by Berger is in (Berger

1989: 185-6), but this later version has deleted both of emphasis and an epistemic tone with respect to the causal determination of the referent.)

Berger explicates here what he calls “F-type terms”, i.e. referential terms whose referents are fixed by perceptual contacts, but later the referential chain is retained by the uses of descriptions also. Even if we ignore the difficulties, over which Berger jumps, with the distinction of the semantic and the speaker reference, what Berger says fails at the crucial point. For how could a description, even a definite one, indicate merely *by itself* in any sufficient way the historical chain, or chains, which originated from the intended referent? A description can do that but only by the trivial way of being part of the chain. This, however, does not enable us to find out anything referentially relevant about the causal chain. (The description could do that only when it says something of the course of that very chain it is part of, but such descriptions are very rare.) On the contrary it seems that the only way a description can indicate referential relevance is that it gives us information about the (intended) referent; but then the historical chain just drops out of the picture. The utterance “Socrates was snub-nosed” contains information about the famous Greek philosopher that he was snub-nosed and presumably this information has been carried via the referential chains to us since antiquity. But about the chains themselves the utterance, hence the description, does not indicate anything. And what it indicates about its referent it does by telling us something about him. It should also be noted that talking about “indicating” the whole chain is perforce a strong idealization. Or at least it is advisable to take it so because the way of indicating is left unexplained by Berger. We do have access to the very latest, the most proximal parts, of the referential chains (in a relative sense, for written notes and other such sources can be old). We do not have at our disposal any stronger means to know the referent than the information the descriptions contain.⁸ I take it that this is so obvious that it does not need any further argumentation. But at the same time it reveals the indispensable epistemic and cognitive elements functioning in the reference by proper names, as well as by the definite descriptions.

A clear manifestation of this is that people typically ask for more information about the referents when they do not have it or when what information they have is insufficient for determined reference. It seems to me that the direct-causal theorists have gone grossly too far in their denials of the referential relevance of knowledge and information in their attempts to avoid anything that would speak in favour of the Fregean sense theory (and the description theories) of reference. Moreover being able to use a proper name referringly on the mere basis of the intentions to co-refer confuses the basic function of proper names - the function possessed by all referring expressions - as referring devices with the uses of them in which reference is *not* taken nearly so easy an accomplishment as the direct-causal theorists would have it via the causal connections. It is very remarkable that the direct-causal theorists overlook that mass of evidence that clearly shows that we are most certain of successful reference only when we have individuating information about the referents. It is easy to claim that the causal connections do all the work, but that is tantamount to fail to see that they go together with the information that does the real work and which the causal links only transmit. Take away that information and people become very unsure about what they are talking about. (Take away, that is, all of the information they have or have had access to, so that no deferential moves are reasonable.)

Why then have the direct-causal theorists lapsed time and again into thinking that the intentions to co-refer retain the referent? As a partial answer I suggest that they have in a way confused things at the theoretical level. It certainly seems that a description indicates a historical chain if one is already looking through the lenses of the direct-causal inclinations (and has accepted, or is sympathetic to, the anti-sense and anti-description theoretic arguments and views). Consequently it is easy to lapse into thinking that, because the theory postulates that there are (idealized) causal chains, we *qua* speakers immersed as parts of them are somehow "in touch" with the original referents. The direct-causal theorists tend to take the causal chains as simply given, even though they recognize that they are communicatively built. But quite ironically this has led them to ignore the very factors that in those very referential exchanges by proper names preserve the referents: information about the referents given by the descriptions (expressive of the senses). Descriptions in use do what Berger says they do: they tell us that the referent is such and such, and in that way they do indicate the referent. But that is how the referential communication typically functions when it is successful and retains the referents of proper names. The picture the direct-causal theorists paint is just too abstract and idealized to be able to get the real explanatory factors into the view – in fact it hides them from view.

The argument just given suffices to show that mere intentions to co-refer are too weak to guarantee actual co-reference between the speaker and the hearer. Conversely, due to the lack of individuating information they could neither repair a damaged link or chain, i.e. they could not by themselves prevent an unintentional reference change. Recall again the well-known case of "Madagascar". This name refers now to the large island southeast of the African mainland but originally it referred to a part of that mainland near it. A natural explanation of the reference change is that new information was brought into the discussion, or that too sparse information was around so that the speakers were prone to confuse the original referent. Once the unintentionally incorrectly borrowed referent became widely accepted as the referent of "Madagascar", the change was robust. Still all the participants to the discussions may have had sincere intentions to co-refer to what "Madagascar" has always referred to. I think that this case admirably underlines a crucial point of comparison between the direct-causal and the Fregean sense theory. It follows as a prediction from the latter theory that the mere intentions to co-refer do not suffice, but information about the intended referent must be involved. The intentions to co-refer are only derivative at best, or subsequent upon receiving information about the referent. It is a typical feature of the language use that when the hearer does not know (in advance) about whom or what the speaker is talking about, or the hearer could not infer that from the discursive context, he demands identifying information from the speaker. Only after receiving that information will the hearer intend to co-refer with the proper name (or whatever referential expression is relevant). Without this *informational check* the cases like Madagascar are very likely to result. And conversely the informational checks keep the referents preserved, not the however sincere intentions to co-refer. (Sometimes the way to referential hell is paved with good intentions.)

Although the Madagascar case is indisputable it seems to me that the whole lesson from it has not yet been drawn. Note first that the emergence of the new causal links, i.e. those based on the island, is via Devittian groundings. In other words the direct-causal theorists could insist that the island is the proper referent because it was named first time in some occasion even if only inadvertently or because of a misconception.

And this is equivalent to giving up the claim that the original object named is the sole referent.

But even that modification is not explanatorily sufficient. For the real problem is how the change of reference became known to have taken place. No one could have been able to realize the change if the causal links by themselves were the only effective factor available. But that was not so: a definite change of reference was detected. The answer to the problem is that the reference change became known because it was realized that to what “Madagascar” now refers differs from what it referred to earlier. And to find that out the mere presence of the past causal links *qua* causal could not have sufficed. So the change must have become realized because people had access to the causal links by *cognitive means*. That is, we have to have such information about the referents that enables us to realize that there is some discrepancy with the current uses compared to the earlier uses. These cognitive means enable us to get access to the causal links but only informationally, only via what is said about the referents. (Except the demonstrative situations where the causal links are perceptually obvious in most cases.) And the natural explanation for *that* is that the causal links and connections only carry information; by themselves they do not have any other role. We are so to say cognitively tuned to information in all cognitive tasks, some of which are referential language uses.

Here we begin to see that the direct-causal theory involves an error of misplacement of its explanatory means. The causal links *qua* causal do not explain reference, for they *could not* do that because of the way we are built cognitively. Only the information that the links carry explains matters referential. (In section 4.2.5 I propose an explanation of the intuition why the causal links – even when communicatively installed, as they should as the more perceptive of the direct-causal theorists have observed – seem so robust mechanisms of reference.)

A corollary of the argument just given is that the notion of reference borrowing needs to be informationally enriched. When a hearer catches upon a new name, a new *fixing* and grounding of the referent takes place. But in order to avoid an unintentional reference change there has to be sufficiently individuating information transfer between the source and the receiver. So we see that reference borrowings with information transmission can be recurrent reference fixings. The direct-causal theorists are obviously right when they point out that the information, in the form of descriptions, can occasionally be false of the referent of the proper name without this severing the fixing link. But they have overgeneralized this observation, and this has led to ignoring the crucial facts effective in the referential communication. For if the transmitted information is misinformation, is totally false of the referent, the danger becomes imminent that we will not be able to retain the original referent. And that happened in the Madagascar case.

I suspect that the insistence by the direct-causal theorists on the mere causal chains in the reference borrowings is nothing but the result of the restricted picture the counterexamples to the Fregean description theory, understood as only a stringent theory of the meaning of proper names, has forced on the direct-causal theorists. The communicative facts manifested in the referential uses of language do not support this picture. In most cases when we use a proper name we talk about and gossip how NN looks, what he has done, his character and acquaintances, and so on. I would even

claim that proper names could not even function as proper names, i.e. naming devices, if they had not this informational task tied to them. That the fixing descriptions may be false or inapplicable in some cases is of course a platitude, but it should not be forgotten that very much more often they are not.⁹ It is this way that the referents are retained and the referential causal nets are weaved through information-transmitting communication.

Devitt seems to grasp the crucial role of information in the reference retention when he speaks of the *acquired abilities* to refer upon catching a new proper name. I suppose that these abilities are supposed to explain our intentions to co-refer, for if we had not any such abilities we could not be able to refer at all. The reason I say that he seems to grasp the crucial role is because he does not say anything sufficiently detailed about the abilities acquired which enable one to borrow reference and use competently the name in question. But perhaps the following quotation helps, to begin with.

"As materialists we expect the advance of science, particularly neurophysiology, to show us that it [the ability to refer, to designate] is a certain sort of state of the central nervous system. It is a state which is brought about in a language user by perception of a naming ceremony...and which is apt to produce (in part) certain sorts of utterances, viz. utterances using the name in question. It is such states, whatever they are, that largely constitute the links between names and objects."
(Devitt 1974: 186)

In that paper Devitt also observes that the various abilities to refer with the various proper names require other abilities. To gain a particular ability to refer with a particular proper name one must already possess the general ability to refer, to use referential expressions to refer (Devitt 1974: 185, n.5). I think that this is correct. One can support it for instance by the Donnellanian notion of referentially used descriptions, i.e. not just proper names. Devitt also suggests, again correctly to my mind, that the abilities to refer can be related also to the non-linguistic behavior, especially to pointing (Devitt 1974: 186, n.8).

As we saw above by themselves the intentions to co-refer are too weak to enable us to retain the original referents. Neither do the mere causal links enable one to do that for the reason already argued: there is no guarantee that the causal links have not been broken earlier. My argument is then that the referential abilities Devitt postulates must be informational to fulfill the role of referent retention - for what other means could there be once the causal factors fail? The abilities must include a sufficiently detailed conception of the referent of the proper name in question.

Devitt relates the abilities to the hearer, but their symmetrical nature is implicit in the case, for obviously the hearer is a future speaker using the acquired name. Therefore the abilities become speaker's abilities to refer as well. And clearly the first speaker with the name must possess the ability to refer. So it follows that the abilities to refer are transitive: via the acquiring of a particular proper name, via the reference borrowing, the same ability so to say travels from one speaker to the next and further on. Devitt acknowledges that information is carried through the communicative exchanges, for he says that "As a result of many remarks using the name type 'John', we acquire many beliefs concerning various people of that name. The beliefs concerning different people are, in some sense, 'stored' separately with their respective

abilities." (Devitt 1981: 34-5) This amounts to the fact that successful reference requires "having referent in mind" (Devitt 1974: 188; Devitt 1981). However, despite that Devitt sees the close cognitive connection from the beliefs about the referent of a proper name to the ability to use that name to refer, he insists that having in mind is best construed as causal relation between the speaker's state of mind and the object, this relation conferring to the speaker the ability to refer to that object.¹⁰

But that just is not so. The reason is, as we observe again, the correct insight of the sense theory of reference (and Devitt himself admits it being such), that of having the referent in mind, loses its explanatory point when it is buried in the causal chains. Devitt accounts for the abilities to refer by their causal inheritance, i.e. borrowing, beginning from the referent which they were originally grounded (Devitt 1981: 34). But with this the relation of explanatory dependence becomes turned around. It is not denied that there are causal relations between objects and the expressions that refer to them. But they *become* referential only because of the information they carry; information from which we choose with respect to the earlier acquired information and in light of which we judge the reliability of the new information as to the existence of the referent. In this connection it is telling that Devitt does admit that the descriptions allow us to borrow an ability to refer also (Devitt 1981: 38). In short, Devitt as well as the other direct-causal theorists confound the "raw" mechanism of reference, the causal link, with the very individuating factor (the sense of a proper name) without which the mechanism could not work, by losing sight of the latter.¹¹ (Again no doubt because of the fixed idea that the Fregean description theory has been shown to be false.)

If one is allowed a harsh manner of argumentation, one could point out that Devitt's causal conception of the notion of having in mind has in fact nothing to do with what having in mind really is, even minimally. For it is characteristic of *that* notion that one can have a thing in mind without any relevant causal connections to or from it. I may have in mind, say, the thickest spruce in the center of the forest outside my window without having ever been inside that forest earlier. Just as Devitt commits a category mistake when he proposes that the causal connections are taken as non-Fregean *senses*, he commits similar one here. The diagnosis can only be that these are desperate attempts to save the causal approach in the face of the grave difficulties that the focus on mere causal links brings with it. True, it belongs to the scientific enterprise, and even maybe more to its philosophical areas, to reinterpret and tear down old and common concepts and distinctions. But here we have a conception that is too deeply entrenched because of the way we are cognitively built. Even if the boundary between cognitive and non-cognitive is not strict (there may be a continuum), the causal interpretation of the senses and having in mind is so far at the other end that Devitt's causal senses just will not do. They so to say try to tie the two ends together at the same time giving up the other end, the cognitive one. This is totally contrary to the view of the senses as semantic-cum-cognitive entities. Consequently I just refuse to take seriously Devitt's merely causal interpretation of the notion of having in mind.¹²

After all this argumentation what answer can we give to Kripke's and Devitt's arguments against the sense theory of reference? The Feynman case from the non-necessity of the senses receives the following answer. To begin with it is to be noticed that the argument itself is not quite adequate. The reason is that the sense expressed by the description "famous physicist" is not an individuating Fregean sense of a proper

name. It is rather a concept expression in Frege's terms. I think that this point becomes obvious in the following quotation: "This is in full accord with the criteria I gave – that the singular definite article always indicates an object, whereas the indefinite article accompanies a concept word." (Frege 1997c: 184) Therefore the referent of "the famous physicist" is the concept *famous physicist*, not the individual Richard Feynman (or any other individual physicist separately, called "Feynman" or not). The description expresses the sense "famous physicist", but that is the sense of the concept *famous physicist* the extension of which is the class of famous physicists. The sense does not even purport to be an individuating sense of an individual referent. Accordingly the "Fregean description theory" could not be faulted by Feynman case.

But if we let this inaccuracy pass the argument boils down to Kripke's causal picture of the links from the baptism of Feynman to the mouth of the man in the street. But I argued that neither the causal links themselves nor the intentions to co-refer suffice to guarantee the original referent. Moreover Kripke overrules the speaker reference in favour of the semantic reference. But that just is the issue: the Fregean theory of reference can be taken as theory that explains the speaker reference as the primary form of reference. The semantic reference is derivative; usually it arises from the individual speaker references through the "informational negotiations", i.e. through the normal information transmission utilized in the language use with proper names. I take the speaker reference to be primary; the semantic references are just speaker reference writ large. That is, it is commonly known who the referent is because speaker information about that referent is commonly shared, and known to be shared through particular communicative exchanges. Sometimes the semantic reference may be incorrect but the speaker reference is not when one speaker knows the real referent but all others believe incorrectly some other to be that referent. Of course it also happens that sometimes the semantic referent is the correct one and the speaker referent is misguided. But in most cases they indeed do seem to coincide; referential use of language could not work efficiently if they did not.

Another way to put my view is that the speaker reference is thinker reference: the referent is that individual to whom one thinks he is referring via the thoughts he takes to be about that referent, i.e. what the sense he grasps individuates. Contrary to what often seems to be assumed the semantic reference is then *not* a property of the linguistic expression in question. The semantic reference is in this sense just a mass phenomenon: there is semantic reference, a community-wide reference, when information about the referent is shared and is known so to be. Therefore one's speaker reference coheres, or does not as the case may be, with the semantic reference.

In the Feynman case this means that the man in the street in the example *does not* refer uniquely to Richard Feynman, the famous physicist. This is the judgement the Fregean sense theory implies. So because the mere causal links are not to be trusted, as I argued, Kripke's Feynman example can only stand if it invokes a tacit application of *deference* thereby giving substance to the semantic perspective from which Kripke amasses his argument. The intuition on which Kripke bases his claim that the man does nevertheless refer to Richard Feynman is not scrutinized sufficiently at all. And moreover in that role Kripke's intuition seems to be the result of the community-wide information about Feynman, for so much is known about Feynman that it is in fact impossible for Kripke, or for any other academic to have escaped being "contaminated" by parts of that information about Feynman. To even to begin to show

his case a direct-causal theorist would need to present a “null case”, i.e. give an example of a proper name about the bearer of which nothing non-trivial is known. But then the claim that the proper name in question is used determinately to refer becomes a slender one; or that is what I predict will happen if queried of ordinary speakers.

What about the Gödel case? It challenges the sufficiency of the role of the senses in reference with proper names. Here again the rivalry between the semantic and the speaker reference is decided in favour of the former notion without any cogent argument, but only by invoking a referential intuition that that is what we would say, or that is how we feel about the proper referent of the name "Gödel". Kripke says that "we simply are not" referring to Schmidt by "Gödel" (Kripke 1972: 294). But in similar terms the case is informationally contaminated to begin with. I want to maintain that if the sense the speaker has associated with "Gödel", expressed by the description "the discoverer of the incompleteness of arithmetics", when Schmidt is the discoverer of the incompleteness, "Gödel" refers in his idiolect to Schmidt. No doubt "Gödel" refers to Kurt Gödel in the idiolects of those who know more about Gödel (the more so if some of them happened to belong to the *very* small class of his acquaintances). That is the semantic reference of "Gödel". That the semantic and the speaker references are in genuine competition here would be revealed if some of Gödel's acquaintances were to meet the innocent and they started to talk about mathematical logic. Presumably it would soon dawn upon that what the innocent has been talking about, when using "Gödel", has not been what the acquaintances have been talking about (namely Kurt Gödel). (Of course the innocent would most likely be corrected and thereby taken in to the network of the semantic reference of "Gödel".) In other words Kripke's example is ruined from the very beginning, just as his Feynman case is. The intuition that tells that we all simply are referring to Gödel by "Gödel", and to Feynman by "Feynman", could not be trusted because of this defect. Were we all as informationally poor as the speakers in the examples, it would become very uncertain indeed that by "Gödel" and "Feynman", with the associated senses, we would be all referring to Kurt Gödel and Richard Feynman. (The same verdict goes with Devitt's two first arguments because they are in effect the same as Kripke's.)

To remind the reader, remember that the counter "the causal links and the intentions to co-refer guarantee the right referent" will not do anymore, for that counter was argued down above. Neither succeeds the counter that I am not entitled to informational considerations because proper names do not refer by the senses. There are two reasons. First, as I argued above, information *does* have a central role to play in individuating referents. Second, the counter is really based on the conception that because the modal and epistemological arguments are sound the senses are not relevant to the issue. But those arguments fail, so this conception becomes irrelevant. As I already pointed out the corollary from all this is that the only way Kripke and Devitt would have succeeded in making a plausible argumentative case would have been by constructing an informational "null case" so that the proper name would not be informationally contaminated from the start. That way the examples, and the use of intuitions, would not be open to the charge of in effect begging the question. (It seems that the Homer case is a better example, and close to a null case, than the Gödel example. But then, as I argued, the direct-causal theory receives no support from Homer.)

What is however adequate in Kripke's and Devitt's examples is that they show that proper names can be used *differentially*. The very distinction of the semantic and the

speaker reference is motivated in part by that distinction. The essence of deferentiality (and not just with respect to proper names) is trust. That is, in using a proper name about the bearer of which one knows nothing nontrivial, or too little by way of the unique individuation of that bearer (as with the description "famous physicist"), one typically associates with it a belief that it genuinely refers because it is in current use in one's linguistic community (or subcommunity). One trusts the other speakers being competent, speakers who know what they are talking about. I would like to propose that this metasupposition associated with proper names, and referring expressions in general, manifests in part the *function* of proper names, i.e. their being used as *naming devices*, singling out certain individuals ("tagging") with the help of the related information (=senses expressed by descriptions). This *metasupposition* is in full force practically in all deferential language use. But deferential practice could not be used as an argument against the "Fregean description theory" for the simple reason that that theory is entitled to include the deferential practice as well. So the postulation of deferential practice, and "name-producers" versus "name-consumers", is no invention by the direct-causal theorists.

However I should like to point out that the deferential practice with respect to proper names may not be that extensive as the direct-causal theorists invariably like to make it seem. That is, are we really as deferential as the direct-causal lore would have us to be? That issue is straightforwardly empirical and so a potential object of study. To my experience – by listening to conversations – we use (at least purportedly) identifying descriptions, and invoke such information, to a surprising extent with the uses of proper names. In a way we are somewhat loath to be deferential: we strongly tend to need to know what we are talking about. That is why we ask questions like "Who is she?", "What are you talking about?": in short we yearn for identifying information. (That is not to deny that there are cases of "pure" deferentiality, but as said my suggestion is that they are minority compared to the claim of their extensiveness by of the direct-causalists.)

In his book on metarepresentations Recanati explicates deference and uses also the notion of trust explicitly (Recanati 2000: 273). I accept his account in outlines. However I do not agree with Recanati's statement that every language has a "deferential operator" related to it as a semantical factor. That seems not to be an unavoidable explanatory assumption. The notion of language related deferential operator should not be taken too literally. Rather I would say that there is a pragmatic factor (in the sense of wide pragmatics, see note 22) concerning communicative comprehensibility such that when one encounters an expression one does not comprehend, or comprehends only partially, one trusts that it has a sufficiently determinate sense which at least some of the other speakers in one's linguistic community possess (the experts most likely) and to whom one can rely on, hence defer. So this supposition is a (wide) pragmatic guide, not a semantic one. And it seems to me to be also arguable that the pragmatic supposition, hence deference, manifests a particular form of our more general *epistemic* tendency to rely on those whom we know, or believe, to know more than ourselves about the specific matters under consideration.

Another theorist that should be mentioned in this connection is Frank Jackson. There is clearly an affinity between his and my views on the referential workings of proper names by the non-experts (Jackson 2004: 270-3). How close the affinity is would

require a further study, but the following has to do here. Jackson talks about “representational properties” related to expressions, whereas I emphasize the underlying “formal” or pragmatic modes of their use here. Still it is quite evident in Jackson’s presentation that the representational properties are about the general features of the contents of those expressions pertaining to the deferential features of them that the competent speakers understand (and can most likely state explicitly if asked to).

The other side of the deferential coin has already been described: if one is not content with merely deferring, upon encountering a new proper name, one asks the speaker about the referent. One wants to become cognitively and epistemically competent speaker with respect to that name as are the others on whom one trusts. I think that it is deferentiality, based on the cognitive and epistemic trust, that also explains the division of the linguistic labour, the resorting to the experts as to the referents of and knowledge about particular terms. The division of the linguistic labour has usually been related to the natural kind terms, but it is obvious that it is a feature of linguistic practice in general, *a fortiori* it holds good with the deferential uses of proper names.

That one is deferring does not entail that one is always referring vicariously by deference, as it were automatically by the references the experts make. It is not the case that the deferring speaker’s use of a proper name is, to put it picturesquely, absorbed into the domain of those who are able to refer identifyingly with it, if the *informational* connections to the latter are too indirect (or that there are no such at all). I just could not bring myself to accept that if there is no informational connection to the ignorant speaker it still is the case that the epistemic state of the community as whole he belongs to enables his deferrings to be genuine referrals. (And even here the right kind of the informational connection has to be spelled out in general.) That the deferring speaker does not refer automatically can be seen when those on whom he trusts do not in fact refer to anything either, as is the case when unbeknownst to them the purported referent does not exist. Conversely the references of those who can refer identifyingly do not automatically “radiate” back to the deferring speaker. Putnam, who introduced the division of the linguistic labour, formulates the issue just in this mysterious way, as if the knowledge the experts possess somehow subsumes also the innocent speakers, enabling the latter to refer because of the epistemic state of the linguistic community “as a whole” due to the knowledge possessed by the experts. He writes that “the way of recognizing possessed by the ‘expert’ speakers is also, through them, possessed by the collective linguistic body, even though it is not possessed by each individual member of the body...” (Putnam 1987c: 228), and “...individual psychological state *certainly* does not fix the extension; it is only the sociolinguistic state of the collective linguistic body to which the speaker belongs that fixes the extension.” (Putnam 1987c: 229). But to my mind this is simply too general and too idealized in that it blurs the sound distinction between knowing nothing individuating and knowing something individuating. In other words, in his attempts to avoid the sense theoretic view and to defend the causal view Putnam – like the other direct-causal theorists - has gone too far in trying to retain a robust notion of reference - too robust in fact. It seems to me that if we go all the way with Putnam’s view, we are led to the position in which we could not help but maintain that we all do refer by all the referring terms we use in our linguistic community, as long as there is one expert with respect to the particular term used. I surmise that this view would turn out to be contrary to facts, were we to ask the other speakers in our respective linguistic

communities at large. It also seems to me that it is very unlikely that the speakers would take themselves to be such vicariously referential omnipotents. The case is instead that at most the innocent speakers, with respect to many terms - especially the esoteric scientific natural kind terms - would say that they “sort of” refer to the bearers of those terms, but not really and that is because they know so little about their extensions. Ditto for proper names.

The deferential supposition does not differentiate between the sense and the direct-causal theory of reference, however. This follows from it being pragmatic and not a semantic factor. It could be that others are competent, and one’s trust on them is reliable because they have identifying knowledge about the referent, so that it is practically certain that the referent exists. Or they may be competent with the name because there is an uncut, reliable, information transferring connection from the referent to the current uses of the name. (Normally when there is an epistemically reliable relation there is also a causal connection; indeed in most cases they coincide extensionally, or so I will argue a little later.) One can say that the deferencing speaker with respect to a particular proper name is temporarily cognitively and epistemically blind. That is precisely why he has to trust the other speakers. The blindness is total when he possesses no (nontrivial) information about the referent neither knows anything about the causal routes of the name (except the last proximal part via which he acquired the name).

The Fregeans can accept deference if the reservations from the informational contacts are taken into account. But even then deference is sometimes only a form of vicarious reference in that it is not quite genuine reference. The reason is that its success depends on there being knowledgeable speaker with respect to the proper name in question. The speaker reference is not fully genuine in such cases, but it is not to be absorbed to the side of the semantic reference either. Only when the ignorant speaker becomes to know identifying information about the referent do the speaker and the semantic references coincide (provided that that information is sufficiently widely known). So we see that the Fregean reading of the notions of deference and speaker reference are more constrained than the direct-causal theory reading. And that I take to be an explanatory virtue.

We still have to deal with Devitt’s third argument, i.e. those cases when a description, or descriptions (expressing the sense) does or do not apply to or fit to anyone or it is false, but the proper name in question still refers. I deal with that argument in section 4.2.5 for it can be put in a better argumentative service there on behalf of my account. Here I just explicate the run of the argument. The claim of the direct-causal theory is that there is a genuine reference made by a proper name even when the descriptions all fail to be true of that referent. (No one of them applies; all misdescribe.) To the extent that the claim leans on deferentiality it carries no argumentative weight against the “Fregean description theory”, as we just saw. But what about if there are no experts to whom to defer? (Here the Fregean and the description theorists’ verdict is that the proper name does not refer at all.) To the extent the claim, now without the role of the experts, leans on the intact causal links it carries no explanatory weight, either; that was also argued above. So what is left with that argument? Only that it seems to have an intuitive feeling about it. And what that feel is based on will be explained in section 4.2.5, thus to further disarm the direct-causal theory. But before that a word or two about the meanings of proper names is in order.

The meanings of proper names

I have been talking quite extensively about information and information transmission through the referential uses of language. The neo-Fregeans have also put that notion in good use, see for example (Evans 1982) and (Luntley 1999). But how does information specifically relate to my Fregean view? At this point, as a prelude to chapters 6, 7 and 8, I need to provide a preliminary answer to this question. I will get to the relevance of information in a roundabout way by first discussing the meaning of proper names. Of course this might feel at least as a minor shock to some readers, for the consensus view has quite a long been that proper names *do not* have meanings. But let us see about that now.

I argued that contrary to what the direct-causal theorists have presupposed the Fregean theory has nothing to do with strong semantic demands like providing proper names with descriptive synonymies. This is the basic conception the direct-causal theorists have of the “traditional theories” of meaning. Just because they argued that the Fregean - and the description - theory fail to give such meaning equivalences they were led to the view that proper names do not have meanings.

There is also another way the direct-causal theorists have argued for their view, a way that adheres closely to the Millian tenet inherent in their theory. According to that argument proper names differ from the general nouns in that the latter have connotations (=meanings) whereas the former do not. For instance the term “chair” means, roughly, a portable object, natural but usually artificial, primarily used for sitting. Proper names do not have meanings in any comparable sense; they just are “non-connotative appellations”, tags. Not even does the name “Dartmouth” mean “the mouth of Dart” but is just the name of a town. (Though that it is located at the mouth of the river Dart may have played a role in fixing the town as the referent of “Dartmouth”. That is, I am talking about Dartmouth 50 21 N, 3 35 W, not Dartmouth, Canada, 44 40 N, 63 30 W, let alone the Ivy League university.)

The weakness of this line of arguing is that synonymies and analyticities (and such semantic notions) do not exhaust the domain of meaning. (If they did, even according to the direct-causal theorists own lights, they would not have proposed meaning theories for the natural kind terms like that of Putnam’s.) It should be noted that there is a general meaning theory for proper names that to me seems more of a degenerate descendant from the synonymy view than a substantial meaning theory. In the formal mode it says that the meaning of a proper name “NN” is “the individual called ‘NN’”. This is intersubjectively sufficiently uniform, as is easy to see. But as I will argue in the next section that view spells out but only the *function* of the proper name as naming devices.

To mention it once again, Frege pointed out explicitly in his Aristotle note – and what he said he did not modify later on – that people have different senses related to the same proper name. To me at least this is tantamount to the denial of any intersubjectively uniform synonymies between the ordinary proper names and the definite descriptions. But by no means does it follow that on the Fregean view the ordinary proper names do not have meanings. Quite the contrary: I propose that the Fregean senses constitute those meanings. The crucial question that should be asked – and should have been asked back in the 1960s - is how to draw the line between those

senses, or sense elements, that contribute to the meaning of a proper name and those that do not. Or are all senses relevant possessed by an individual speaker?

The answer to the latter question is no. We have already encountered the reason for that: when a proper name has, for an individual speaker, only a general concept related to it, that name does not refer individually. That was the lesson we learned from the Feynman example: the sense expressed by “famous physicist” has as its extension the class of famous physicists only. Hence only identificatory senses are relevant. (However those senses can be expressed by structurally more complex concept expressions, like “the pupil of Plato”, when they have an unsaturated functional element - “the pupil of _” – and the “saturating” element – “Plato” together in one overall structure.) But that is all that is required of them. So when the descriptive senses one relates to the name “Aristotle” are, for instance, “the most famous pupil of Plato”, “the author of *Poetica* and *De Anima*” and “the teacher of Alexander the Great”, these express the meaning of “Aristotle” for that speaker. Contrary to the direct-causal theorists view there is no relevant difference to how the general terms like common nouns and predicates pick up their extensions, except that in the latter cases there is an additional step involved in the Fregean view. That is the step from the sense of a concept word to the concept itself, and then to its extension consisting of the objects characterized by the concept. But the referential structure is the same, in both cases ending up with objects. The only difference is the ranges of the objects: the general terms refer to classes of individuals, i.e. they refer collectively (or distributively); proper names refer to individuals separately. But that does not suffice for any arguments to the effect that proper names do not have meanings.

The meanings of the general terms can then be taken to be individuating as well. They individuate the classes of objects (but also substances like water). The classes consist of individuals that are in their turn individuated by more focusing meanings (=senses). There is nothing qualitatively, let alone metaphysically, different in the manner the meanings of the general terms relate to things from the manner the meanings of proper names relate. The meaning of a general term differentiates its extension from the extensions of the other general terms; the meaning of a proper name differentiates its referent from the referents of the other proper names.

Now we can turn to the role of information transmission in communicative exchanges. Speakers use descriptively expressed information related to proper names in question when they talk about the happenings and whereabouts of other individuals. The information these descriptions transmit becomes to form the senses, or elements of them, of proper names. Some senses are immediately individuating, “one-shot” senses, like the one expressed by the description “the fifth president of Finland”. Of course it must be realized that the individuating effect is relative to what the speaker knows at that moment: if one does not know what president is, or what Finland is, it is unlikely that the description speaker-refers uniquely to Risto Ryti – or for the audience.) The senses of some proper names become individuating only in piecemeal fashion. For instance none of the following descriptions make a unique reference to an individual but together they do: “was nearly killed at Mommila in November 1917”, “was the head of the board of the Bank of Finland”, “was the president of Finland”.

Linguistically speaking proper names have meanings and cognitively speaking they have senses. That is, proper names have meanings because they have senses. And

those senses are what the speaker “grasps” when he acquires information about the individuals to whom the names refer. The grasping of most sense elements need not be conscious. All in all, then, I conclude that the claims to the effect that proper names do not have meanings, but only refer and tag, are based on a misconception of what meanings are in general and what the senses are in particular. It seems to me that the claim that proper names have no meanings is the result of a myopic view. By this I mean that those who have so claimed have been unable to take a step back and look at the issue afresh. For I surmise that if they had done that, it would not have been likely that it had escaped their notice that they have in fact mistook the *multiplicity* of the senses related to proper names for the lack of their meanings. This myopia is understandable to some extent because the meanings of the general terms – common nouns and predicates – are relatively easy to explicate, at least to the extent that their characterizing features can be agreed on by most speakers. (That does not mean that they possess necessary and sufficient conditions – as indeed they seem not to.) But just because a particular proper name usually has quite a many pieces of information related to it, hence multiplicity of the sense elements, it is all too easy to lose sight of that fact and be led to think that proper names do not have any meanings at all.

That proper names do have meanings, not a single uniformly shared piece of information descriptively expressible, is still denied by the direct-causal theorists. As I already mentioned for example Scott Soames writes that there are no uniform meanings (or senses) that proper names have, hence they are not synonymous to any definite descriptions: “However, there is no substantial descriptive content associated with [simple proper names – like “Soames”] that remains constant across all contexts...” (Soames 2002: 26); see also (Soames 2005) in many places). I think that it is a high time that we turn this direct-causal tide and accept that proper names *do* have meanings, or contents, in a perfectly natural sense of that notion that does not lay any demand on the meaning being constant and shared by all competent speakers. Therefore we could do well to discard all synonymy demands of them also.

It also seems to me that there is no qualitative semantical difference between proper names and the common nouns when looked at from the point of view according to which meaning is what relates to an expression in such a way that enables us to use it in communication. According to that view we have general characterizations of, say, chairs and humans, but like with the latter class of entities – and their names – we could (begin to) name chairs and relate their particular properties like colours and shapes to those individual names. That would obviously facilitate the individuating and identificatory aspects of our “chair talk”. (Think of a hypothetical sect that worships chairs as gods.) Us humans being what we are, then, it would be no wonder at all that the meanings of proper names contain much specific information; chairs are rather static entities and their uses are few. In other words those who have thought that semantically the common nouns differ from proper names have confused the level of extension-characterizing information with some principled semantic category.

Note, by the way, that uncovering that confusion gives further support to my view that the referential use of language is a form of human action, and indicates that it involves a ground for a generalization (the substance of which I will not develop here): semantics of the linguistic expressions is subordinate to the uses of those expressions. And because these uses require information meaning requires information. Therefore the particular things in the extension of the common nouns are used the way they are;

it is indifferent to us in most cases what particular individual belonging to the extension in question relates to the utterances (a chair is a chair is a chair). The semantics of those expressions, i.e. their meanings, is left at the general level of characterization.

The general change of perspective I suggest here is based on the following observation. Proper names and the general terms have been taken to be semantically different classes of expressions because they have been focused on from the linguistic perspective. (But then again, in a way, how else they could have been looked at in the philosophy of language?) But if we adopt the neuroscientific and in general neuronaturalistic perspective, and focus on the linguistic expressions across the board as “tools of the brain”, it is no longer adequate to compartmentalize proper names and the general terms. The reason is that the way the brain works – gathers and utilizes information in commerce with both external and internal reality – allows us to make no explanatory semantical difference between these two classes of expressions. As I argued the only relevant difference is about the *range* of the referents of the respective expressions. From the semantic, *a fortiori* referential point of view that is accidental.

Of course it could be claimed, similarly to Kripke but forgetting the synonymy demand, that if the meanings of proper names are just the pieces of information about the bearers, the interest of the Fregean sense as well as the description theory diminishes too much. This claim needs to be answered, not because it could be correct but because it brings the background of the issue out in the open. The situation changes 180°: the reasonable sense and description theorists have not ever demanded anything stronger than that the senses and the descriptions function as the factors which single out the referents. In particular no uniform, “constant content”, related to any proper name has been claimed to exist. So the diagnosis is that the direct-causal theorists have mistakenly thought that they have been occupying the reasonable ground all along – the ground in which the sense and the description theorists have in fact been standing even before them. Only by believing (and propagating) the view that the sense and the description theorists have been committed to the theory of the meaning of proper names in some semantically strong sense like providing synonymies (“constant content”) have the direct-causalists’ arguments been able to gain credibility. But that credibility should vanish the very moment we realize that it has been based on an utter straw man as to the *meaning* theory of proper names allegedly held by the sense and the description theorists. The general argument is that to infer from the lack of synonymy to the non-existence of the meanings of proper names is *non sequitur*. The inference would go through only if synonymy (and such strong, maybe even too idealized, notions) was the sole semantic notion involved (or even relevant). But it is not; only the too narrow theoretical and methodological climate back in the 1960s has made it look so.

It also seems to me that the *multiplicity* of the pieces of information is, in effect at least, the same thing as Recanati talks about in his theory and by which he explains the direct referentiality of proper names and the feature REF (Recanati 1993). This observation is supported also by the tendency to relate direct reference of proper names with the “no meaning” view. However Recanati argues for that view by stating that proper names have *linguistic* meanings; this is what I called above the “degenerate descendant view” from the synonymy demands; that the meaning of a proper name “NN” is “the individual called ‘NN’” (or some equivalent to that). This may be an

acceptable middle position between the stringent synonymy view and the Fregean view because it is about the linguistic meaning: the Fregean view is primarily about the cognitive meaning. (And Recanati is a neo-Fregean in that he postulates “psychological” modes of presentations also.)

That proper names have meanings but that they are accessed mostly tacitly in use and so need not even be descriptively articulatable in all cases is strongly indicated by those proper names that have become to “live their own lives”. These names include the names of the famous figures like “Einstein”. It is rather common to hear it said that someone is “a little Einstein” or, maybe more often, “you are no Einstein”. What the former utterance is understood to be saying (or implicated or whatever – I do not care here for the Gricean distinctions) is that one is bright and imaginative, and the latter that one certainly is not. These uses of “Einstein” would not be comprehensible at all if people did not relate certain properties to that name, based on the character and achievements of Albert Einstein, the physicist and natural philosopher. These properties are what the senses, i.e. the meanings, of proper names express.

There are also two corollaries we can draw from the above observation. The first is that the category of “partially descriptive names” that Scott Soames has argued for (Soames 2002) is not a distinct category of referring expressions on its own right. The descriptive names – like “Mt. McKinley” and “professor Saul Kripke” – just make explicit the most central piece(s) of information as elements of the meanings of those names. For example “Mt. McKinley” makes it explicit that “McKinley” is – of course – ambiguous but because of that very ambiguousness it has become common practice to disambiguate its uses to refer to the mountain (location at Alaska 63 4N, 151 0W; 6194 meters high). If the “simple” proper names functioned in the common use as directly referential tags there would not be any semantic need to invoke explicitly an identifying descriptive element.

The second corollary is that proper names are *not* only direct naming devices. The reason is the tacit (descriptively expressible) meaning elements. In fact this is rather obvious when something is said about someone by using a proper name. As an example: “He’s a regular Koba”: “To say of someone ‘he’s a regular Koba’ was to accuse him of impropriety, foulmouthedness and lack of moral refinement.” (Hingley 1974: 39) Or consider the shout I heard recently during a heated debate “don’t Bush it!” I take it that the shouter warned the other to remain within the sphere of decent argumentation. Obviously these utterances and many like them make sense only because proper names used have such information related to them that is commonly known to characterize the intended bearers of the names. So proper names *can* be used also *predicatively* and not only referentially, contrary to the view in the direct-causalist camp.¹³

It seems to me also that, as a corollary to the second corollary, Recanati’s account of direct referentiality as being a *type* of reference (see section 3.4) is not sound. Recanati argues that proper names are type-referential, i.e. that they are not just used referentially (as for example are the descriptions according to Donnellan’s referential-attributive distinction) but *are* nothing but always directly referential as syntactic and semantic types of expressions. But if some proper names are used as predicative expressions, it immediately follows that proper names are not type-referential as such. They are only as token-referential as the descriptions used referentially. That proper

names function in almost every case of their use as referring devices should not be mistook for the lack of the possibility of predicative use. So the referentiality of proper names is only a *statistical* phenomenon, not the sole function of them. (In the next section I will also argue that the feature REF that Recanati relates to the directly referential terms is neither the sole feature of proper names.)

Proper names and the basic schema of reference

I have been arguing that information carried by the communicative causal chains, not the chains themselves, is what guarantees the retention of the referents of proper names. As I mentioned among the neo-Fregeans Gareth Evans already focused on the role of information. But he discussed mainly demonstratives and indexicals. In general the neo-Fregeans including later writers like Recanati (1993) and Michael Luntley (1999) have not scrutinized the nature of information at any deeper level. They have been content with coarse taxonomies like perceptual information. So my approach here – and in the following chapters – is an attempt to improve on the usual way of thinking about information relevant to reference.

In light of the essential role of information being the decisive factor in reference the direct-causal theorists could change their argumentative gear and suggest the following. They may admit that the Fregean theory, in its modern informational version, does explain how we use proper names *in communication*, i.e. locally. But the causal story explains why there is the *continuity* in those uses, i.e. why the Fregean factors are able to accomplish their referential task. But this suggestion will not stand closer scrutiny. To show why that is so I want to radicalize ontologically the argument from “no causal links but only information”. I argue that there *are* no causal chains *qua* causal, but only *qua* information transfers. Therefore the causal chains the direct-causal theory postulates are even powerless as explanatory factors. The argument is quite straightforward. What the direct-causal theorists have been talking about are in fact informational relations *through and through*. To understand this let us take a general look at what really takes place in so-called causal interactions. When an event is causally preceding another, i.e. when we talk about something causing something other to happen, we are in fact talking about *information transfers* (or information chains, if you like). The talk about causal relations is only a folk idea about contact relations that taxonomizes events in contrast to others. For instance a billiard ball contacting another due to which event the second ball moves is said to be a causal event. But what happens is that (part of) the momentum of the first ball is transferred to the second. Or in biochemical process calcium ions inside a neuron get into contact with the protein kinase molecules and that contact changes the conformation of the latter, which in turn changes their ranges of activity. Of this process we say that the second messengers caused a cascade of changes that eventually leads to gene expression. When a sentence follows from another, as when one applies *modus ponens*, that relation is logical; still the information the sentences express drives the inference. And when you say "Aljechin was the most creative chess player so far" you transmit the information your remark expresses (i.e. your thought). If I know who Aljechin was, if I am able to identify him, hence know to whom you referred, I may agree with you. If I do not know who Aljechin was, presumably I ask you. Still, information has been transmitted. So instead of plain causal relations we should be talking about the different kinds of relations of information transfer, at least in the appropriate philosophical and scientific contexts.

It is true that we do not yet know much about the structure and dynamics of information and how they constitute the informational interactions. Neither is this the place to engage in that study, the “metaphysics of information”. But as a modest hypothesis I want to propose that information is what emerges in the interactions of the different structures of matter and energy at all levels of reality. Change the surroundings of the type of interaction and you might have an interaction with different information involved (like when the same syntactical expression has a different meaning in different contexts). And it might be that these “levels” are themselves only relatively robust dynamical structures of matter-energy, i.e. maybe reality *is* information through and through. Everything that is, is in unceasing bath of information, inside out and *vice versa*. But information is not substance; rather it is intertwined processes. My argument is then that the *causal connections as such do not explain anything*. It is the information transmitted, and utilized in the cognitive contexts, that counts explanatorily. (In some cases it may even be that the medium is the message; the alluded biochemical processes seem to be this kind of information transmissions. Anyway, even if the medium-message distinction is ubiquitous it does not affect my argument either; at the level of reference only the message counts, not the medium by which it is carried.)

It is rather trivial that the causal links have been selected in the folkish conception as the installations of the mechanism of reference. For in the great majority of cases there are intact causal links. But that is because the information is smoothly transmitted; because of that we *know* whom we are talking about in those cases. (Of course we add to the information by associating pieces of other information, and by inferring new pieces from the acquired ones.) The essential point is that the causal links, if such there are *qua* causal, at most only carry the information. Therefore it is very easy to understand why the misplacement of the real explanatory factor of reference has taken place: we have a deep entrenched image of causal connections – like lines connecting things - and we are not used to think that the simple connections, by whatever physical media, really *consist* of the varieties of information transmitted. (We do not perceive the information *an sich*. But it is there, and our brains pick it up and utilize it.)

My general argument is then that the direct-causal theorists could not have recourse to the claim that we need not know anything about the causal links for there still to be genuine reference. They could not do that because there are no causal links *qua* causal; it is all information transmission. And that requires – I would even be willing to suggest that it *nomologically* requires – that for information to be functional it must be registered *and* be tuned to, i.e. the receiving end must be able to utilize it. (If there were no “utilization” of it, there would be no effects.) In the case of the referential use of language that tuning on is the ability – on the basis of already acquired information and inferences – to pick up individuated information about the referents of proper names.

The direct-causal theorists are likely to object. They can ask "does not this radical picture on information transfer succumb to the argument that, with respect to the sense theory of reference, it could not as such differentiate between information and misinformation?". That is, my proposed view of the information transfer, instead of the causal relations as such, does not explain the fact that in the communicative exchanges there are occasionally transfers of misinformation in which no true (non trivial)

information is given (or implied). Is not one forced to maintain in regard of such cases that because no correct piece of information helps to retain the referent, it must be (even if only what we call folkishly) the causal connections that only count referentially in retaining the referent? According to this argument the causal connections, so to speak, are revealed as referential through the fog of the misinformation they carry. In other words, do not the causal connections nevertheless seem sufficiently robust by themselves? If that were so the direct-causal theorists could then conclude that there are referential uses of proper names, the associated information of which is false but which still do refer. Hence the sense and the description theories fail after all. Here is then the crucial version of Devitt's third argument from the inapplicable descriptions but still referring name, the version that I mentioned quickly earlier. If my arguments up till now are correct, this is the last straw for the direct-causal theorist to clutch. But it also sinks as I will argue at length in the rest of this section.

The explanation the direct-causal theory gives for this referential claim is that there is a causal chain of the uses of the proper name originating from the naming of the referent. But I argued above that the causal communicative exchanges, even fortified with the intentions to co-refer, do not suffice to retain the referents. So we seem to be back at the level of the intuitive support to the direct-causal theory. At that level the core question that never seems to have been asked is why does the causal connection seem so strong a factor that it is claimed to overrun the (mis)informational impacts associated with proper names? Why does this intuition seem to have so strong a grip on most of the direct-causal theorists in the way that it is let to discard the at least as strong empirical facts showing the informationally driven referential communication with proper names, facts which do not support the direct-causal theory?¹⁴

One must realize that the clash here is a symptom of a deep underlying difference between the sense and the direct-causal theory. When the issue is viewed from the dialectical angle the direct-causal theorists adopt, the causal connection easily seems compelling over the senses as the explanation of the mechanism of reference. But when it is viewed from the angle of the actual communicative exchanges the sense theory overruns the direct-causal considerations: in most cases we do not have access to the causal history of the name; we could only use the information associated with the names. The reason is simple: the past causal connections are out of our current control. And as we saw they are out of our control even when guided by the intentions to co-refer: there is no purely causal guarantee that the referent would not have changed. In other words a successful reference borrowing needs support from the informational factors. (To object that the direct-causal theory does not require us to have an access to the causal links is correct but helps not a whit because, as I argued, it leaves the retention of the original referent as *the* referent unexplained.)

In personal communication Panu Raatikainen has pointed out to me that the causal links being out of our control is also Kripke's view. However I am not claiming that it is not. My argument is that because we rarely do know the past causal route, the claim that those links constitute reference is much too weak as an explanatory claim; indeed it is much too unconstrained even with the intentions to co-refer. The intentions bring needed cognitive respectability to the direct-causal theory but that is still explanatorily much too weak. With this meager cognitive resource we are unable to assure ourselves that the communicative causal links are preserved intact from the naming to the

current uses of the name. Taken purely causally a proper name might refer or it might not for all we know, and there is nothing we can do about it. (Except if we can control the causal route at the very communicative situation with the presence of the referent.) A quick argument that shows the danger of the uncontrollability of the past causal links is that there may be cases in which we use proper names exactly the same way as usual but without knowing that they do not refer to anything. In such cases the causal links have become cut or the individuals we purportedly refer to do not even exist. That such cases are rare, for all we know, speaks not for the causal but for the informational retention of the referents. (As an example consider the postulated planet Vulcan to explain the anomalous Newtonian orbit of Mercury.)

Put in another way the charge is that the notion of reference given to us by the direct-causal theory is too idealized, and too simplistic, for it to be able to explain the referential practices. The *fact* is that in the actual practice of the referential language use we eventually do require cognitive control over the referents of proper names. That is, even if we tend to defer we trust, and demand if need be, that some users of proper names do know what and whom they are talking about, i.e. are able to individuate the referents sufficiently well. The names have also an epistemic role in that the representations of their (purported) referents *qua* referents are situated within our beliefs. (These are sometimes called “notional worlds”.) Epistemically speaking it is the overall coherence and cohesion that counts with the uses of proper names. The direct-causal theory just fails to account for that.

But let us suppose both that information is discarded and that the causal-cum-communicative factors do not suffice by themselves for the retention of the referents - what then will we find underlying the referential intuition invoked by the direct-causal theorists? It seems to me that in addition to the causal picture there is something having to do with proper names themselves from which the intuitive referential force of the names supposedly derives. This feature shows itself in those cases when there is no referent but a proper name is used as if there is such, i.e. when we do not (yet) know that there is no referent of the name in question. For the sake of the argument we can ignore the fact that even these uses of proper names take place with respect to an informational background, with respect to the beliefs purportedly about the referent and other things related to it. So all there is left in this type of case is the proper name itself and that it has been used as a genuinely referential expression.

Some theorists have considered the view that the *meaning* of a proper name, "NN", is best explicated as "there is an individual called, or named, 'NN'" (Bach 1987; Recanati 1993).¹⁵ This view seems to hit on some important property of proper names; it might even partly explain what is meant when it is said that proper names serve as mere tags ("Proper names can perhaps be regarded as associated with the referential feature *being named or commonly called 'NN'.*" (Evans 1982: 312, n.10)). But I propose that instead of explicating (only) the general linguistic meaning of proper names, it is far more fruitful to take the explication as an indication of the basic factor underlying all referential expressions. In developing this theme we can take as our starting point Recanati's argumentation for the general meaning view with respect to proper names. Recanati postulates a property of directly referential singular terms he dubs REF:

"...referential terms have the following property: their meaning includes a special feature, which I dubbed "REF", by virtue of which they indicate that there is an object the referent of the term..." (Recanati 1989: 241)

He also relates REF to thoughts and what he calls *de re* concepts of individuals (as the referents of proper names):

"But, in another sense, it is possible to think about objects directly, thanks to the special nature of *de re* concepts. They are such that *the thought in which they occur is intended to characterize the reference itself independently of its satisfying the very concept which is used to think of it*. This property... is the mental counterpart of the feature REF in language. It is achieved by means of a simple architectural trick: putting our conceptions of an object in an informational file supposed to derive from that object makes it the case that a thought containing a pointer to that file is about the object itself." (Recanati 1993: 130)

I would like to adopt the feature REF but transform it from the philosophical surface phenomenology of the referential expressions into a scientifically argued postulation belonging to the domains of developmental cognitive psychology and neurolinguistics. From this perspective it becomes clear that REF is indeed an indication of the basic function of the referring expressions, an indication of a developmentally fundamental mechanism functional in language learning.

The studies of Jean Piaget, as well as many later ones, have shown that the sensorimotor patterns of action begin to develop in the newborn infants from the very first encounters with their surroundings.¹⁶ It is arguable that this whole patterning is based in part on an innate neural specialization in service of the social behavior and the establishment of being in rapport with the entities in one's surroundings (Brothers 1990). Here I am interested in the relations of these patterns to language learning. I condense a long developmental story to a very short space, ostensive and demonstrative pointings shown to an infant by a caretaker, and then made by the infant itself, are not unproblematic. In other words learning takes time: the infant has to integrate the different factors that become to constitute the abilities to refer demonstratively. He has to learn from rudimentary gestures, gazes, touches and the caretaker's vocalizations which of them go together in the intermodal perceptual and motor patterns with respect to which external objects. Later proper names form a special class of vocalizations (or so we may assume) in that most of the objects, of which they are vocalizations, are other humans like the caretakers, and that everyone has one's own name (in contrast to general concept names like "chair"). It is likely that only after some time the infant will be able to comprehend the function of the names as naming devices from utterances like "this is Gottlob" and "Bertrand is sleeping". Then the infant begins to comprehend proper names as symbols that are used to refer to those certain individuals that they are the names of. This is one form of the birth of the symbol function. By that time the child has also learned to distinguish proper names from the common nouns, so he is able to comprehend the difference between proper names and the descriptions attributing properties to the bearers of the names.

There is evidence that pointing has acquired referential function after about 7 months (Butterworth & Glover 1988; Messier & Collis 1996: 450). Recall that Devitt hypothesized that pointings are related to the referential abilities. Now we see that he

is basically correct when the relevance of the pointings is explained from the perspective of the developmental cognitive psychology. After pointing has become a robust function in the child's cognitive tool kit (in the beginning of his "linguistic turn") there comes first the single-word utterances, then two-word utterances and so on to the fully structured sentences of one's home language (or languages, as it might be). The learning of demonstratives and personal pronouns as well as indexicals as pointers to the surroundings takes time also. When the names of the pointed-out and reached-for objects get linked to the information received from them (constituting the sensorimotor and cognitive-perceptual information) the basic abilities (Devitt's general abilities) of the referential use of language are in place. The specific individual abilities to refer that Devitt postulates can be explained as formed from the cognitive-perceptual information one acquires upon encounters with the referents. (Later the connections become also indirect, as in reading about the referents). The connections of the specific abilities to the general ability to refer are cognitive: in hearing and in reading the name in question can engage the information related to it (its neural representation), and in thinking and in speaking that information (or parts of it) engages the name and its relation to the neural structure realizing the general ability to refer, thus producing the referential use of the name in question. (More about that in chapters 7 and 8.)

I dare to propose that once the general and specific abilities have been acquired, with respect to proper names, the descriptions, demonstratives and indexicals, by then the child has acquired *basic schema of reference*. It is a functional neurocognitive and linguistic structure that enables us to deal with our surroundings, first literally by guiding hands and gazes and then linguistic expressions. Later when the child acquires more and more information about the world, the linguistic devices assume more and more the role of referring devices within the child's inner epistemic world; he hypothesizes, uses counterfactuals, engages in fiction and pretend-play with them.

The basic schema of reference is not primarily a structural neural entity, a type-identifiable structure. But as a structure it is *functional* one. It is the functionality which counts out the type identity, for, trivially, the blind, the deaf and the mute are able to use language referentially (natural or signed), but the relevant respective areas in their brains that would support the type-identification of the basic schema are lesioned as compared to normal speakers. Although I can not go into the details of the neural underpinnings here, the following must be said. Presumably the basic schema of reference consists of many networks (and networks of networks) from the primary receiving sensory areas to the secondary and the tertiary multimodal and associative regions, together with the subcortical nuclei, most likely thalamus, hippocampus and basal ganglia. Also Broca's area seems to be involved for there is currently evidence that it partakes in ordering relations. As it happens additional support for the idea of the basic schema of reference can be found in the results that implicate a role for Broca's area in the recognition of movements and action patterns and in their executions (Rizzolatti & Arbib 1998). A very interesting is also the suggestion made by Rizzolatti and Arbib that Broca's area is a counterpart, a homolog in evolutionary terms, of the area F5 in the premotor cortex in monkeys. It is an area phylogenetically preceding the speech area in humans that also participates in other communicational functions. In other words the basic schema of reference could be a species-wide phylogenetic functional structure - humans only have succeeded in exploiting it massively with language. Broca's area seems to be the place where it "all comes

together", i.e. where the build up of the perceptual and cognitive information associated to the representations of individuals relate to the linguistic expressions and their utterings, these being the end products of the neurocognitive processes active in the particular occasions of the referential use of language.

The philosophical point of the postulation about the basic schema of reference is that it explains the intuitive referential pull that we saw above proper names to possess. I further propose that this intuitive pull is also what Recanati's feature REF indicates, and what Salmon is after with his notion of reference *simpliciter* (though he misconstrues it). All these rather fleetingly explicated notions owe their existence to the sensorimotor and the cognitive-perceptual patterns which develop during the frequent ostensive and demonstrative encounters with one's surroundings as an infant.¹⁷ Now the important argumentative point of all this is that the referential intuitions (and REF, reference *simpliciter*) are not primarily specific to proper names at all. They are the results of the more basic neurofunctional structure that makes possible our capacity to use all kinds of linguistic expressions to acquire specific referential function. Consequently the counterexamples to the Fregean theory given by the direct-causal theorists, in which the referential intuitions are called for help, have been misunderstood. Contrary to what is widely believed they do not point to primary semantical-cum-referential features of proper names, in contradistinction to the definite descriptions, but to the *preconditions of any referential use of language*. The basic schema does not favour proper names over the descriptions, or the demonstratives over the descriptions as (directly) referential devices. Without the development of the basic schema of reference no linguistic expressions could refer at all.¹⁸

One can maintain, following Recanati's account, that there is still a certain difference between proper names and the descriptions in that the former are clearly more directly referential than the latter. But it seems to me that what the persisting difference indicates is the rather obvious difference between proper names and the descriptions: the former *name* individuals whereas the latter do not (except when, as is sometimes said, "they have grown capital letters" like "the Holy Roman Empire"). The descriptions do not name but, well, descriptively characterize their referents. But notice that at the level of the modes of presentation proper names and the descriptions are on a par. As Recanati himself argues the mode of presentation of a proper name "NN" is the name "NN" itself with the meaning "the *x* called 'NN'". The mode of presentation of a description is, of course, the description itself (expressing part of the sense of the associated name). In other words, because in most cases the descriptions express *parts* of the senses of proper names, it is not to be wondered at that there is a difference between them with respect to their referential directness, i.e. between naming objects and specifying their properties. So if we were able to explain the difference between naming and specifying the referent, even when the referents of proper names can only be identified by their senses (parts of) which are expressed by the descriptions, we would have been able to account for our proposed extension of the feature REF. However I have to postpone that explanation for a while.

So with the basic schema of reference the first intuition about the "intrinsic" referential force of proper names is answered. But what about the second intuition, the very causal picture which is supposed to display the referential power of the causal

connections from the referents to the later uses of their names? How can it be explained away?

Let us begin with the causal picture itself. The Thales example can again be taken as the representative example of that picture. What is wrong with it is that it has been magnified to such a proportion that it has become to support a much too an idealized picture. One just could not help getting that impression from the example itself. There is the original referent of the name "Thales" and he is given that role because of the causal connections to the later uses of "Thales". There just is no room for error in this picture; the causal connections are supposed to be drawn and preserved intact. (If there were an error, for example a "cut", the whole example would lose its relevance.) Now what are such situations of the referential language use where there are no errors, or at least the possibilities for errors are minimal? It needs only a quick reflection to see those are the demonstrative and the ostensive situations. These are the primary situations in which a speaker singles out his intended referent to his audience, and the audience comprehends (mostly by perceiving) what or who the intended referent is.

To put this in another way I claim that the picture of reference which the direct-causal theorists have drawn is simply a generalization of demonstrative reference, and very idealized at that. The demonstrative references are those situations in which the speaker perceives the referent in his immediate vicinity - just as it is in those situations when the basic schema of reference develops. In reality the causal routes are spatiotemporally extended, hence at the mercy of the vagaries of all historical processes. But what is crucial in the direct-causal picture is that all the factors that are postulated to guarantee the determination of the referent are compressed into it, just as they are in the demonstrative situations. And just like in those situations, in the layout of the Thales example one takes himself, as the speaker, to be within it, even if within a short reflective distance. This tacit move is easily understood to have taken place because the fundamental mode of our referential *Dasein* is to be always "situated" in some particular surroundings of objects, persons and events. In the direct-causal picture one, so to speak, looks omnisciently at reference from the demonstrative perspective, i.e. as if one partakes in it as just in another demonstrative referential situation.

I think that it adds force to this diagnosis when we take into account that it is sometimes claimed, on behalf of the primacy of the causal connections, that the referential case is parallel (i.e. stronger than analogous) to the perceptual case. In the latter case the causal connections are taken for granted as providing the most adequate explanation of perceiving something. For example Evans writes

"The absurdity in supposing that the denotation of our contemporary use of the name 'Aristotle' could be some unknown (n.b) item whose doings are causally isolated from our body of information is strictly parallel to the absurdity in supposing that one might be seeing something one has no causal contact with solely upon the ground that there is a splendid match between object and visual impression." (Evans 1985a: 13)

But it seems to me that it is precisely this "absurdity" that remains a real possibility because the parallel fails: presumably it would be absurd to maintain that someone was or did something if there were no causal contact whatsoever, but that just does not

extend to the language use in its entirety. The reason is that language is also stimulus independent vehicle of storing and transmitting information. That is, language use, especially referential use, enables us to talk about things to which we may not have any such causal connections that would amount to closeness to the object on which the causal theory of perception relies, and which Evans presupposes in the above quotation. Mere causal connections will not do for reference for the simple reason that there are too many of them, most of which presumably are fleeting and irrelevant even by the causal theorists' own lights. The parallel Evans draws manifests again the tacit assimilation of the demonstrative situations and the presupposition of the general adequacy of the referential causal relations.

But, as we have been seeing in this chapter, the vagaries in the causal routes of proper names destroy the idealization inherent in the causal picture; there are no causal guarantees against cut routes, mixed routes, meshed routes, diverging routes, and other such contingencies. The intentions to keep on co-referring will not do for the referent retention; for that informational impacts are required. The mistake of the direct-causal theorists is the age-old and familiar one discussed in the philosophy of science of causation. For instance, a match is lit and causes a fire. But it would not have caused the fire if no oxygen had been present, no appropriate materials had been present, the weather had been too windy, and other such conditions. In the referential case the sole operative cause has been taken to be the use of a proper name and the surrounding cognitive and informational conditions have been largely overlooked. But just as with the match causing the fire the surrounding conditions are at least as much needed for the referential effect. I would go as far as to state that, in both cases (and in such cases in general, of course) the surrounding conditions are the *enabling causes*. A robust referential route for a proper name is possible only because it transmits *information* about the referent.¹⁹

It is understandable that the demonstrative situations have tacitly become the model for the direct-causal theorists: as I said our everyday *Dasein* is immediately contextual and we do not just speak about the things in our surroundings, we use and manipulate them as well. The causal connections because they are short and perceptually sharp, manifest themselves in these activities; touching and using objects is a form of demonstration. The causal events are "short" in that we usually manage fairly quickly to achieve the effects we intended; they are "sharp" in that at the perceptual level unexpected and loose effects like misperceptions are not common. The cause-effect events are typically quite isolated at the perceptual and practical level. So I propose that the natural explanation of that is the contribution of the basic schema of reference. For us that is just as well because it strongly indicates that behind the intuitions and the picture of the direct-causal reference there is in fact a single explanatory factor. The basic schema of reference, so to say, keeps the direct-causal theorists tacitly on its grip.²⁰

What comes to the Thales case, which has prompted the whole argument that the causal routes guarantee the reference of a proper name even when not a single piece of true information is known about the referent, the Fregean sense theoretic prediction prevails: when no true information is available we just *do not* refer by that name to anyone, or to anything, neither from the speaker's reference perspective nor from the semantic reference perspective. If it is later discovered that there *is* a referent for that name, then only *from that point on* we would be referring to the bearer of that name,

but not before. In this case it even suffices that we only know that there is a referent; we need not know any specific descriptions of it, for we may have become to know that there is the referent by a source, or several sources, which only tell that the name "NN" was (or is) the name of NN, that NN was, indeed, called "NN". Or we may be able to infer from some other pieces of information that there was that NN. (Or even that there was NN, who we call "NN" but who was back then not so called.) These sources may even deny explicitly that anything else is known about NN. Both of these types of knowledge are still sufficiently individuating in that we know that "NN" is a name of a certain individual. The reason is that the information the source conveys (or which we infer) relates NN, at least implicitly, to other pieces of our knowledge about the times and places which help to confer the epistemic reliability on the source. But if that "surrounding story" is lacking we do not accept the claim that there was NN but that we just know nothing about it (else than that it was called "NN"). That we do not accept that claim seems to be a common fact of our referential practices, but I also intend it to be adopted as a constraint on that practice: because the Fregean sense theory seems to be correct in informational terms, we should adjust our judgements with respect to such cases which have seemed problematic up till now and deny any purported references with a proper name when nothing is known about the purported referent. This shows clearly how the reference of proper names is deeply cognitive and epistemic matter. The causal connections are there but they *are* the information transmitting links from the sources to us and they are *constituted* by the cognitive activities of the speakers at the occasion they put the sources into use, i.e. understand what the sources say or imply about NN.

So I urge that no matter how compelling the direct-causal intuitions may feel to so many, they should be shrugged off, because they are misinterpreted and much too idealized even then. They indicate nothing with respect to the referent retention of proper names. The demonstrative picture underlying the direct-causal theory and the referential intuitions are just those general but deep-entrenched, mostly tacit, conceptions through which we become used to perceive and manipulate things. Instead I propose that we take a fresher look at what really takes place in communicative exchanges. That should help to leave aside the direct-causal picture of reference and the related intuitions.

To set the record straight we should note that the confusion is not restricted only to the direct-causal theorists. As we saw above even some neo-Fregeans are victims. About a possible case in which all information about a purported referent is misinformation, maybe incorrect rumors about the referent of "NN", Evans writes that

"Nevertheless [the speakers] have got hold of rumours and claims *about a particular man*. As it was used in conveying the misinformation to them, and as it is used by them in further transmitting that misinformation, the name has, and is understood by the consumers [of the name] to have, a quite definite reference, provided for it by the practice of those who know some individual as NN (namely the producers [of the name])." (Evans 1982: 385)

As I argued above it may be the case that "NN" could be used as if it were a proper name of a particular referent even if there never was any individual named "NN". And if the whole information is misinformation, on what grounds one is supposed to hold firmly that there ever was anyone to whom, or what, "NN" purports to refer. (One

could not even state, as Evans does, that there is a particular *man*; for all that is known, i.e. nothing, there might have been, say, a particular dog named "NN".) If the producers of the name who know something about the referent of "NN" are still alive and around Evans' point is correct, but only because it invokes *deference* of the later name-consumers to the former name-producers. But as I argued earlier that feature can not be used against the sense theoretic account. It seems to me that Evans confuses here the pragmatic metasupposition I proposed in the previous section. That goes with all referring expressions in the deferential uses, with the functioning of proper names at the object level as devices referring to a particular individual named "NN".

We may eventually discover that the rumours are all false of a certain NN. But that by itself does *not* prove that "NN" was used to successfully refer to NN in the middle period, when all of the original and correct information came to be replaced by the rumours and till the discovery of the rumours. Just like the direct-causal theorists Evans has here succumbed in front of the force of the referential function of proper names and taken it to be the essential referential substance of them as such. But that is not so; it is only a misplaced intuition. Moreover, just like the direct-causal theorists Evans presuppose that once a proper name acquires a definite referent (by the producers) it thereby stays with the proper name everafter till it "dies", i.e. drops out of circulation. That is, it is presupposed that the reference of a proper name could not cease and be re-established. I myself do not see anything in this presupposition such that would compel us to accept it. On the contrary it seems quite possible that particular referential relations can cease to exist and be re-established. This is so because reference is epistemic and cognitive matter. To repeat, reference is not causal but informational affair; hence when information related to a proper name is complete misinformation, reference ceases.

Now I have reached the end point of my counterargument to the semantical argument. The main result is that the sense theoretic approach has become vindicated. In short, the informational factors, which find their way into the senses of proper names and which are expressed by the definite descriptions (though not all of them could be so expressed) are the only factors that can explain the reference of proper names. Ontologically the (what we have used to call) causal factors are informational. The demonstrative and the ostensive references are no exception, for the informational perceptual and cognitive factors determine the referents in those situations. The direct-causal theory is correct in these situations; it only fails to notice that the factors it relies on are very Fregean ones even there.

The Twin Earth argument

Senses *do* determine referents

As we recall the thrust of Salmon's Twin Earth argument is that the conceptual contents of proper names are not the same as their information values. The argument proceeds from the assumption of the conceptual sameness (and the sameness of the respective mental states individuated by content) of the *Doppelgängers* to the difference of the referents of their (syntactically) same utterances ("Hubert weighs 165 pounds"), because earth Hubert weighs 165 pounds but Twin Hubert 165.000000001 pounds. According to Salmon this shows that the information values of "Hubert" differ in the two utterances without the conceptual content of the utterances differing.

The first impression one is likely to get from this argument is that it is much too an easy one. For who would deny that in the envisaged situation, in the two different environments, it is obvious that the *Doppelgängers* refer to two different things? And more to the point, there is no evidence gathered from Frege's writings that would imply that he would have denied this had he thought about this kind of possibility (i.e., put in somewhat anachronistic terms, were he to have thought about possible worlds). But we should still ask does it follow from his theory of the sense and reference that he should have denied it. That is, is the Twin Earth argument a good one against the Fregean sense theory of reference after all? (I let it also pass that strictly speaking this type of scenario begs the question against the counterpart metaphysics of individuals. That is, Earth Hubert and Twin Hubert are counterparts of each other, *strong counterparts* one could say, because they are identical except the minutest difference in weight.)

The answer is negative, though it takes a few steps to show this. Let me begin with Salmon's notion of information value that he equates with the referents of the linguistic expressions that encode the information (objects, properties and relations). Information value is, then, closely related to truth, as can be seen in his version of the Twin Earth argument: the statement "Hubert weighs 165 pounds" is true in Earth but false in Twin Earth. What in Frege's theory corresponds to this? Frege makes a distinction between grasping a thought and making a judgement. To judge something is to assert that a thought a sentence expresses is true or false. To grasp a thought is to entertain a thought without any claims in regard of its truth-value. The grasping of a thought corresponds to entertaining the conceptual content of the mental states of the *Doppelgängers* in the Twin Earth argument. Both *Doppelgängers* hold that Hubert weighs 165 pounds, so there is no relevant difference in that respect in spite of the fact that the sentence "Hubert weighs 165 pounds" is true in Earth and false in Twin Earth. By hypothesis they also grasp the same thought that the sentence expresses. Consequently the thought as the sense of the sentence "Hubert weighs 165 pounds" should determine the same referent, but it does not. How could we exempt Frege's theory from this failure (so to say merely 10^{-9} failure)?

It should be realized that Salmon's talk of the conceptual content seems intended to be understood in psychological sense, for he says that conceptual content is subjective. Strictly speaking this alone is sufficient to charge Salmon's argument as defective, or even being beside the point, because it is put in terms that are clearly un-Fregean: Frege denied categorically that psychological states have anything to do with semantic and cognitive matters, except that grasping a sense is a psychological act. The senses and the thoughts are third realm denizens and therefore intersubjectively objective in Frege's framework, but mental states are only subjective occurrents. However I apply the Principle of Charity and take Salmon's notion of the conceptual content in the objective or the "third realm" sense. (I am not committing myself to any third realm entities, as should be clear from the fact that I am out to naturalize Frege's theory.)

I said that the first impression one is likely to get from the Twin Earth argument is that it is too easy. This may lead one to suspect that there is something implicitly wrong with it. My diagnosis is that the argument makes an incorrect assumption about what the sense theory is committed to; an assumption that has become accepted as unproblematical, or that it has not even been realized that it has been accepted. It

seems to me that the assumption can be detected with only a little reflection. The manner of the determination of the referent by the sense seems to be the source of the trouble. It seems to be misconceived by Salmon and the direct-causal theorists alike. As I pointed out it is very unlikely that Frege would have denied the conclusion of the Twin Earth argument. Consequently this puts the central explanatory principle of the sense theory of reference only seemingly in jeopardy. But we need to elaborate this a bit more.

Consider what would happen were some of us, except the two Huberts, transported (or teleported?) to Twin Earth, and the Twin Earthians to Earth, without either of us noticing this. This procedure does not affect the senses, the conceptual contents, of our linguistic expressions. This will show up in the respective subsequent behaviors: they will not differ from what they would have been were we not switched. (It is part of the Twin Earth scenarios that the respective Earths are exact replicas of each other at least when it comes to the relevant features of the issue.) The linguistic behaviors will not differ either, so when we will refer to Hubert – i.e. whom we take to be our Earth Hubert - in Twin Earth we will not detect any difference and will not know that it is the Twin Hubert to whom we have referred. In the same vein the Twin Earthians will not detect any differences in Earth. Thus our respective senses do determine the referents just as before. It is only that the truth-conditions and the informational values of the sentences in which "Hubert" surfaces are different. But that is *not known* to anyone. We judge the sentence "Hubert weighs 165 pounds" as true in Earth, which it is, and we judge it true also in Twin Earth, only it is not. But this difference in the truth-values is irrelevant to the thesis that sense determines referent. In other words, the direct-causal theorists assume too strong version of the thesis. The reason is that they have assumed from the outset a non-cognitive and non-epistemic interpretation of the determination relation. To put this point bluntly, it is not and it never was part of the Fregean sense theory of reference that senses would determine referents beyond any reasonable failures, in all possible contexts (or possible worlds). When we will discover and become to know, or only to believe, that there are two Huberts, we will divide the previous uniform referential practices with "Hubert" accordingly. And that follows because upon the discovery we become to differentiate between the two senses of "Hubert". That is, the senses become to include an element that tells that the Huberts are different individuals. Of course we may still believe that Twin Hubert weighs 165 pounds when he in fact does not, but - to repeat - that does not matter for reference determination because we have become to grasp the new senses related to "Hubert" and only these senses are relevant for the judgements and truth-evaluability of any sentences in which "Hubert" occurs from there on.

There has always been a persistent tendency to take the Twin Earth examples from the God's eye point of view, i.e. from the epistemologically omniscient perspective. But strictly speaking this assumption is not warranted when the issue concerns the explanatory adequacy of the Fregean sense theory of reference. That is because the assumption clashes with the central tenet of the sense theory, viz. that senses function primarily as semantic-cum-cognitive factors. This involves that they are intertwined with our (and the Twin Earthians') epistemic practices. When we do not yet know that the two Huberts are two different persons, the situation is just as it should be from both our and the Twin Earthians' mundane cognitive and epistemic perspectives. The accusation that sense does not determine the referent *uniquely* is imposed from the perspective of epistemic omniscience. It is given from within the presumed current

knowledge used in the argument that distinguishes the two Huberts. (One could say that Salmon commits himself to a general use-mention confusion, as do all those who use the Twin Earth scenario to argue against the Fregean senses as reference determiners.)

The sense theory of reference predicts that when the senses are or remain the same, the referents of our referential practices remain the same also. This prediction is born out in the "switch test": both the Earthians and the Twin Earthians will keep on referring exactly to what they take to be the same things in both planets just as they did before the switch. No referential differences are detected. (The switch test is also used by Jerry Fodor but to a different purpose (Fodor 1987).) If any referential deviance from this uniform pattern should take place, the only reasonable source of explanation of it, given the premises behind the Twin Earth argument (and the assumption, which I share, that the senses remain unaffected during the transportation), would be traceable to some difference in the senses, which are the only explanatory cognitive and semantic factors in the theory under consideration. But this would destroy the Twin Earth argument because the sameness of the senses of the *Doppelgänger*s is essential to it.

So we can conclude that senses *do* determine the referents to all semantic and cognitive purposes. It is true that there is not one and unique referent in the case of "Hubert". But that does not matter because the same referents are determined (or should we say cross-determined) *uniformly* by the same senses, as long as the difference is not known. Twin Earth scenario involves a misidentification of the truth-conditions with the "meaning-conditions" of thoughts (and sentences). Moreover the Huberts are *taken* by both the Earthians and the Twin Earthians to be the same referents with respect to the knowledge they have. The knowledge they possess is in its turn explained by its being embodied in the senses and thoughts the people in both communities grasp (which are all same by hypothesis, for otherwise the whole Twin scenario becomes suspect argumentatively). It is the grasping of the thoughts that matters for reference determining for the constituent senses of thoughts, not that we always grasp only true thoughts (let alone that we would know that). This latter option does not follow from the Fregean sense theory, neither is the theory committed to it. As long as we keep on letting the perspective from epistemic omniscience influence our philosophical conceptions of what is going on in Earth and Twin Earth, we will be unable to comprehend the situation from the perspective of the sense theory proper. Only when we discover that there are two Huberts in the two planets, the senses that we grasp when we use the name "Hubert" have become changed and divergent. I propose that a proper Fregean, albeit somewhat metaphorical, way to drive the point home is to suggest that in the latter situation our *range of grasp* of the senses and the thoughts have changed. Strictly speaking in the situations up till and after the discovery that there are two different referents we grasp new senses in addition to the old ones. In this manner our knowledge changes also and that has its effects on our referential practices, just as the sense theory predicts.

Epistemic symmetry principle

Because the point I am making here is rather radical when compared to the received view it may help if we imagine that our current epistemic state of affairs is that of what it was before the discovery that there are two Huberts. In that scenario we (currently,

as it were) use expressions to refer to what we take to be the same things and take the things to be the same (in extension) without being bothered at all by the possibility that they might be distinct (even when in fact they are so but unbeknownst to us). In other words the senses which we currently grasp determine the referents of the expressions they are the senses of. Later scientific research, or just happy everyday coincidences, might reveal the distinctness of the referents. Then we will change our referential practices accordingly with respect to those expressions. The reason is that, to keep on putting it in the Fregean terms, we have become to grasp new senses and thoughts. What we now grasp as being the case and what we now judge as being true or false will also change in the relevant respects. The point is that whatever is the "real" encoded information and the "real" information values of the expressions that make up the sentences we use we may not currently know, so they could not affect our current referential practices either. Except in the way that we may be more guarded in our judgements and referring; that is, we may become more conscious of the *metaprinciple* which says something like "the referent may turn out to be a wrong one, or even non-existent, and the thought to be false, so proceed cautiously". Consequently it is misconceived to argue against the Fregean account by the Twin Earth examples to the effect that the senses do not determine the referents.

This temporal comparison of the collective epistemic states reveals what I like to call, for want of a better term, an "epistemic symmetry principle" (ESP). It is modeled after the basic symmetry principles of physics, for example the equivalence of temporal points, spatial locations and directions. The importance of the physical symmetry principles lies in the fact that they all imply some conserved quantity or some invariance (conservation of energy, momentum and angular momentum in the case of the three above mentioned symmetries). But the symmetries are not restricted to the domain of physics. Different symmetry considerations are relevant in chemistry and molecular biology as well. And from there it is only a short step to neuroscience and cognitive science – or so I propose here. (On the different kinds of symmetries prevalent in nature and in many scientific disciplines, not just physics, see (Mainzer1996).) Applied to the sense theory of reference the analogical invariant is the state of knowledge: when the knowledge of a community does not change, the senses do not change and neither do the thoughts. From this it follows that the referents the senses determine do not change either. Only when the symmetry between the respective epistemic states is broken there results a change in the senses, thus possibly in the references. This is analogous to the case when a physical symmetry is broken. Then one value of a relevant parameter or quantity is singled out in contrast to all the other previously possible. The physical symmetry principles and invariance laws have a powerful role in the construction of physical theories and explanations. They dictate that something could not happen in our universe. For instance you could not construct a *perpetuum mobile* because if that were possible energy would not be conserved (in closed systems like universe). Similarly the epistemic symmetry principle I propose says in effect that one is not referring to a thing as distinct from another if one does not know or believe anything distinctive about it, i.e. one does not know or believe about the distinctness directly, or does not know or believe indirectly something that would allow him to infer the distinctness. I further propose that the epistemic symmetry principle (and there may be others like it) should be accepted as one basic principle of cognitively oriented semantics. (Purely formal semantics are different things altogether, but they have no explanatory relevance to the studies of reference.)

It follows from the epistemic symmetry principle that no semantic and referential comparisons are allowed with respect to the referring expressions in the different contexts if the totality, or relevantly restricted part, of the knowledge of the speakers in those contexts differs respectively. Neither are comparisons allowed across the contexts at a single instant if such knowledge is imputed on the contexts from outside, knowledge to which the subjects of those contexts could not have access, or which did not or could not have had its effects upon them at that moment. From these restrictions it follows that any semantical and referential cross comparisons between Earth and Twin Earth, before and after the discovery that what is called "Hubert" weighs 165 pounds in Earth and in Twin Earth a tiny minuscule bit more, are prohibited if the respective communities could have no access to these facts in a way that would change their range of grasp of the relevant senses. It also follows that *we* are not allowed to impute, and are not justified in smuggling in tacitly, our epistemic perspective to either of the two communities when they are compared at any temporal moment. The Twin Earth scenario violates both of these prohibitions, so it violates the epistemic symmetry principle, hence the Twin Earth scenario could not be used against the Fregean sense theory of reference as a form of cognitive semantics.²¹

It is sometimes argued that the respective behaviors of the *Doppelgänger*s are not the same because the objects toward which, or with which, they behave are different (Recanati 1993: 200-1). For example, an Earthling drinks H₂O and a Twin Earthling drinks XYZ. One could not invoke then the so-called narrow contents alone in the explanations of the respective behaviors. But this argument fails to show its case. This follows from the considerations we just found with the epistemic symmetry principle. The reason is that, as in the situation of the comparison of the referential uses of the (syntactically) same name in the switched contexts, both *Doppelgänger*s would behave in the same way as before were they switched to each others' contexts. Once again, it does not matter in these cases that *we now* know the differences. What matters is what the speakers and thinkers *in those contexts* know and what differences *they* do detect.

It can be accepted that the Twin Earth arguments point to some important feature related to the notion of reference. But the Twin Earth scenarios exaggerate the substance of that feature. For it need not be anything stronger than the (tacit) recognition by all competent speakers that if the uses of two (syntactically the same) proper names, and referring expressions in general, in fact do have two distinct referents, contrary to what has been thought, then their senses should also differ from one another. But that metaprinciple does not by itself result in any differences in the senses of the names in their actual uses as long as there is no indication that they may in fact have different referents. So the Twin Earth scenarios do not support arguments against the sense theory of reference; they only give a vivid explication to one tacit metafeature underlying the notion of the reference.

One is likely to ask why should ESP be elevated to such a dominant role? That is, why should semantic matters mesh with epistemic ones, or with any other kind of matters at all? This is a sound question and deserves an answer, if for no other reason than to let us avoid the charge of begging the question.²² The answer evades also the charge that the principle is *ad hoc*, i.e. proposed only for the purpose of disarming the Twin Earth argument against the sense theory of reference. However the basic justification I can give has to be a short one here. The epistemic principle acquires its motivation from a

larger perspective within which I look at the semantical issues and, of course, propose should be adopted at large. I take it that all parties involved agree that semantics is concerned primarily with the content of the linguistic expressions and their referential properties. But there would not be any linguistic *representations* if they did not express the results of our thoughts and perceptions. I take this Gricean view as being fundamentally correct of the way we humans are cognitively built. Languages are so tightly meshed with our cognitive (and practical) functioning that a semantics which aims to study the contents of the linguistic representations, but which ignores the cognitive and the epistemic matters which manifest themselves in our linguistic communications, does so only to its own peril. This is the reason why the sense theory of reference can gain explanationwise because it relies on the cognitive and the epistemic factors in its explanation of the reference of proper names. The direct-causal theory, on the contrary, loses explanationwise because one of its core claims is that the semantics of the reference of proper names is non-cognitive and non-epistemic matter.

In addition to these general considerations one possible counterargument to ESP needs to be dealt with, however. It may occur to a direct-causal theorist to argue that just because ESP is epistemic *and* it is applied to the Twin Earth scenario, it could not harm the direct-causal theory. For the Twin Earth scenario represents in effect Kripke's notion of qualitatively same epistemic situation. For instance one could be in such a situation with respect to "Phosphorus" and "Hesperus" but the names having different referents. In that situation there are two distinct planets and the two proper names have two distinct referents – whereas the qualitatively same epistemic situation *we* find ourselves in there is only one planet, Venus and the two names has Venus as their sole referent. That is so because "Hesperus" and "Phosphorus" are rigid designators.

That is true but it has not the assumed relevance with respect to ESP. For either the application of Kripke's notion is beside the point, or it is an open issue whether the application of that notion or the application of ESP is primary. It is beside the point when it is realized that the proper domain of the Fregean sense theory of reference is our cognitive and epistemic states at a particular situation. In other words, because reference is cognitive and epistemically constrained phenomenon, metaphysical or such strong considerations are not argumentatively applicable. And if Kripke's notion is taken as applicable, it could be faulted immediately. The reason is that it smuggles in philosophically illegitimately our current epistemic state into the epistemic state built into the example (as it does in the Twin Earth scenario). Only because of that could it be claimed that "Hesperus" and "Phosphorus" are not distinct objects when compared to the epistemically qualitatively same counterfactual situation.

ESP provides also an antidote to the notion of "omniscient observer" that the direct-causal theorists have invoked since Donnellan's explicit postulation (Donnellan 1974) to Joseph Almog's recent flirtation with it (Almog 2005). ESP counters the use of the omniscient perspective (of an Almighty Theoretician) because that notion violates the boundary within which reference research should be confined, i.e. reference as it is manifested in our ordinary referential use of language. These uses are informationally-cum-cognitively based – apart from the preconditions like the metaprinciple uncovered above, which the direct-causal theorists' arguments really only indicate – and the omniscient perspective ignores that empirical situation. And that it does, I surmise, because it also mistakenly sees the causal links as sufficiently powerful by themselves

to be able to retain the referents. That is, the application of the notion of omniscient observer again confuses the demonstrative reference for reference as a whole.

As such the omniscient perspective is rather innocent; all explanatory approaches involve focusing on their domains at varying levels and magnifications. But the application of the omniscient perspective/observer that claims that the finding out, retracing, of the causal links from the current uses of the name back to the original naming event would be available is far from being an adequate claim. For such omniscient retracing is not possible for us as theoreticians – as it were observing the whole causal history (hence Donnellan’s label “causal-historical explanation view”) of a proper name at once (the demonstrative illusion again!). We *can* retrace the links, however, but in limited contexts and with the help of information only. First, in the very naming events the causal links are at sight through perceptual contact; second, by finding out what was, and has been said, by whom and in that way as far back as is possible. But that usually is not far back: we do remember but only few of those events where we have communicated with the proper name in question. (This observation is, then, correct by the direct-causal theorists, but it does not offer a stronger support to their account, not even when the links are preserved via recorded sources like books and conversation notes. For when information runs out, omniscience whimpers away.)

ESP is violated also in the usual arguments the direct(-causal) theorists amassed to show that despite the appearances in the Frege puzzles the speaker does believe that Hesperus is a planet and that Phosphorus is a planet, even when he does not believe that Hesperus is Phosphorus. The reason is that they adopt the theoretician’s omniscient epistemic perspective who knows that Hesperus is Phosphorus is Venus. But this step just makes the singular propositions expressed, according to the direct(-causal) theory, irrelevant to the speaker’s concerns. (Of course the speaker may surmise that it is possible that there is only one self-same planet; but even if he does, but does not believe it, he presumably does not take himself to be expressing the same beliefs with respect to “Hesperus” and “Phosphorus”.) In other words, because of ESP we are able to distinguish referential concerns from the issues pertaining to the truth of those beliefs in relation to the world external to them. Singular propositions are, so to speak, what we may aim at, but because all our semantic practices are also epistemic, singular propositions are relative to the total knowledge possessed by us at a time. The referential and the truth-pertaining issues should be kept distinct because what is the case but not accessed cognitively – believed or known – does not have an effect to the semantic contents of thoughts. That is why the singular propositions as usually conceived of are not taken as cognitive constructs, except when accessed via some mode of presentation of them. But that is what makes the trouble, for the theoretician has access to what is the case through many distinct modes of presentation of the singular proposition. (And singular propositions *could not* be accessed by any other means: reality *could not* be semantically accessed by any other means than *cognitively*.) For example Salmon’s postulation of “guises” reflect this aspect of the singular propositions, i.e. in effect it admits that for the singular propositions to play any explanatory role in semantics they need to be tied to some kind of cognitive access to the referents. That is why the Fregean “intuitions” prevail but the direct(-causal) ones fail when the focus is on the actual practice of referential use of language and thought expression. Singular propositions might be all right for formalist semantics that does not, or need not, honor ESP. But for explanatory semantics they do not suffice.

The direct-causal theorist could still object and state that it is because “Hesperus” and “Phosphorus” are rigid designators that Kripke’s conclusion does follow. But it follows only if the notion of rigid designation stands on its own. However I will argue in sections 4.4 and 4.5 that it does not so stand. Instead it presupposes the Fregean factors in that the referents have to be antecedently individuated. Therefore I take the liberty to conclude already before that argumentation that ESP overrules Kripke’s notion of qualitatively the same epistemic situation in matters referential.

Internalism and senses₁₋₃

Internalism is an approach in the philosophy of mind and language is commonly taken to give no explanatory role to the environment of an organism in the constitution of the semantical properties, the contents, of the mental states of the organism. Sometimes this is expressed by saying that internalism eschews “wide content” and endorses only “narrow content”. Wide content is (a type of) content that involves at least some objects, events, relations and properties external to the organism. In contrast narrow content does not find any sufficient contribution for them in the cognitive and psychological explanations of behavior, *a fortiori* the linguistic behavior including the referential use of language (when these are given in terms of content, that is). Another way that is used to characterize internalism is to say that it is committed to methodological solipsism as a research strategy, i.e. that it does not explain behavior with the help of such mental states that would imply or entail entities other than those states whose content depends only on the internal states of the organism under explanation.

The notion of narrow content and its relation to the notion of wide content can best be explicated by an example. Fortunately the very Twin Earth scenario provides us with such an example. What goes on inside the psychological or “cognitive lifes” of the two *Doppelgänger*s is, by hypothesis, exactly the same. Whatever the Earthling thinks and whatever concepts she possesses, the Twin Earthling thinks and possesses also, and conversely. In the Twin Earth scenario these thoughts and concepts are taken to correspond to the Fregean thoughts and senses. Consequently the respective identical senses of the two *Doppelgänger*s are taken to be the narrow contents that internalism endorses. When Eartha thinks that Hubert weighs 165 pounds, and Twin Eartha thinks that Hubert weighs 165 pounds, the narrow contents of the thoughts are the same. But in contrast the wide contents of the two thoughts differ. The wide content of Eartha’s thought is that the Earth Hubert weighs 165 pounds, and the wide content of the Twin Eartha’s thought is that the Twin Hubert weighs 165 pounds. The wide contents, then, include the external objects of the thoughts. The wide contents are usually interpreted as functions from the narrow contents and the contexts of the use of the expressions to the truth-values. So because the contexts of the two *Doppelgänger*s are different in the Twin Earth scenario, with respect to the referents of “Hubert”, the wide contents are different.

Internalism is also taken to endorse the Fregean thesis that sense determines reference (or equivalently that intension determines extension). In the Twin Earth scenario this obviously implies that narrow content, as the sense of a proper name, determines the referent of that name. In other words, nothing that lies outside of a cognitive agent is needed to account for the individuation of the thoughts of that agent, or for the

sufficient explanation of the agent's use of the language he speaks. All that would be needed for the semantic purposes, as well as for the explanation of behavior, is to be found inside the cognitive economy of the speaker.

It is usually the second clause, that of the sufficiency of the internal factors of an agent for the explanation of the use of the language he speaks, that the externalists have attacked. (So in general the externalists are those who deny the explanatory sufficiency of internalism - but usually also its adequacy on other counts.) If all that it takes to explain one's referential use of language is to be found within oneself, within one's cognitive economy, it follows that the referential behavior of the two *Doppelgängers* should be fully accounted for by only internalist principles. But it is claimed that their respective referential behaviors differ, despite the fact that they are internally exact replicas of each other. So according to the externalists it follows that 1) internalism is false, and 2) the Fregean sense theory of reference is also false, because the senses of the expressions are the same but the referents differ. The moral the externalists draw from the Twin Earth scenario is that the context, the referents or extensions of the terms used by the cognitive agent and also by the other speakers in the agent's linguistic community, are indispensable for the specification of the semantic properties of the expressions of the language.

As we remember from section 3.2 Salmon equated the substance of the Fregean sense theory of reference with the thorough descriptiveness of the senses, in the manner that the senses give the "purely qualitative and/or general properties" of the referent they determine. This characterization means in effect that the senses are purely conceptual representations that do *not* involve anything non-conceptual. In particular the senses are supposed to involve neither anything that is in its turn about other objects nor any contextual factors in determining their referents. This is taken to be equivalent to the internalist view. If the senses are purely conceptual representations, then they could not explain the referential divergencies the Twin Earth scenario is alleged to reveal. Related to the three subsenses that we discussed earlier, the Twin Earth argument is taken to show that sense₁ and sense₂ are different. The reason is that the sense₁ (the narrow content, the conceptual content) is the same for both *Doppelgängers*, but sense₂, the manner the referent is secured, differs because the referents of the name "Hubert" are different. By implication, then, the conclusion from the Twin Earth argument is that there are no uniform Fregean senses.

This combination of the externalist argumentation and taking the Fregean sense theory as an internalist theory has become very widely accepted. But the combination is misconceived and the pattern of the counterargument should be familiar by now. The first thing to notice is that there is nothing in Frege's writings that can be forced, somewhat anachronistically at that, to the mold of internalism. Neither is there anything in his writings that would support the imputation of the doctrine of the senses being purely conceptual representations or representations with purely qualitative and/or general properties of their referents. *The situation is quite the contrary.* To realize this it is sufficient to give a couple of extracts from Frege's corpus.

In concluding this section it should be mentioned that sometimes internalism is taken to have its roots in Cartesian considerations. It is claimed that according to internalism mental states and representations could be exactly the same as they are if the external world did not exist. (Shadows of Descartes' evil demon.) This is the way Putnam

introduced internalism together with methodological solipsism (Putnam 1987c). But that conception is erroneous. Descartes pondered of the possible non-existence of the external world because he was interested in methodological and epistemic issues; he only used ontological or metaphysical examples in his argumentation against wholesale skepticism. Modern internalism and methodological solipsism are only psychological and, derivatively, semantical views; the existence or not of the external world is a red herring. Internalism does not deny (nor is committed to deny) the existence of the external surroundings of the subject, quite the contrary. (Besides there is a discrepancy manifest in Putnam's Twin Earth scenario in relation to his characterization of methodological solipsism in that the relevance of the existence, or not, of the external world with respect to the *Doppelgängers* is not even remotely hinted at. The point of the Twin Earth scenario is the psychological sameness but semantical difference between the *Doppelgängers*.)

Internalism does not claim any kind of psychological and semantical, let alone ontological, independence of mind on its surroundings. It is a doctrine about the representational capacities of mind *in relation* to its surroundings. It states that our mental-cum-representational constitution is epistemically relevant (but not in regard of the problem of skepticism); being surroundings-constrained in the strong externalist sense, and how that shows up in externalist semantics, is too strong epistemically. We just could not let us have such semantics and theory of reference that involves that what we currently have become to know (at separate temporal spans, epistemically speaking) influences the contents of the thoughts and the meanings of the linguistic expressions of other subjects in epistemically different (temporal) situations. That is the point of my postulation of ESP.

The (partly) non-conceptual status of the senses

It is true that Frege did say about the senses much less than we would have liked him to say. But it seems to me undeniable that Frege acknowledged the role of environment, both causal and contextual factors, in the *constitution* of the senses and in their reference determining roles, even if he did not say that in so many words (and in so many places). Let us begin with the notion of pure conceptuality as it is explicated negatively by what Tyler Burge has to say about *non-conceptuality* of the senses. He writes that

"...term 'non-conceptual' does not imply that no concepts or other mental notions enter into a full statement of the [contextual] relation...The crucial point is that the relation not be merely that of the concepts' being *of* the object-concepts that denote or apply to it...perceptual content...illustrates the sort of element independent of semantical or conceptual application that is essential to the notion [of contextuality]." (Burge 1977: 346)

Burge goes here against the assimilation of the Fregean senses, as the modes of presentation of the referents, to the senses as the determiners of the referents. He does that mainly on the grounds that the latter notion requires non-conceptual and contextual relations, and that "they go beyond what a thinker grasps in thought" (Burge 1977: 358).

Another characterization of the conceptuality involving purely qualitative properties is given by Robert Adams. According to him a property is purely qualitative if, and when, "it could be expressed, in a language sufficiently rich, without the aid of such referential devices as proper names, proper adjectives, and verbs (such as 'Leibnizian' and 'pegasizes'), indexical expressions, and referential uses of definite descriptions." (Adams 1979: 9) When this view is combined with Burge's emphasis on the non-conceptual angle on the senses, we get essentially Salmon's account. Note that Adams' view also gives voice to the linguistic version that demands "thorough descriptiorality" of the senses.

All seems to be well as far as one does not ask whether the notion of pure conceptuality has any sufficient plausibility after all. The argument is that if it does not, the whole substance behind Salmon's (and many others') argumentation collapses. In other words we could then apply *modus tollens* instead of accepting the straightforward argument to the effect that so much worse for the Fregean senses. Take Socrates for example. He can be described by invoking qualitative properties as being the most famous philosopher who was sentenced to death and who drank hemlock. Even in this case there seems to be an implicit indication to human kind. But not just that. For if one takes a closer look at the very description, one realizes that it does not enable one to use "Socrates" in one's linguistic community (and in one's thoughts) to the extent that proper names, with their individuating descriptions, are typically used in referential exchanges. I may know that the most famous philosopher drank hemlock, but that by itself does not yet guide me at all in my cognitive and linguistic activities if I do not know in addition other, mostly implicit or tacit facts related to the bearer of the name "Socrates", facts about ancient Greece and Athens, about democracy, history of philosophy, and such things. These facts, hence the senses of the expressions related to them, are not purely conceptual, or at least they involve elements that are not such.

In the famous Aristotle note Frege spells out two different senses of "Aristotle" out of many. In the first is mentioned "the pupil of Plato" and in the second "born in Stagira". This shows that the Fregean senses can contain the senses of other individuals and in this way they can be object dependent. This in no way militates against the *Russellian* demand that all ordinary proper names involved in the descriptions must be further analyzed to their general properties. No Fregean denies that for example "Plato" and "Stagira" have their own senses even when they form part(s) of other senses. But the point is that the Fregean senses need *not* be "analyzed through". Maybe they can but there is no obligation for that in Frege's account. (And Frege never demanded that.) We see here, again, how the careless assimilation of Frege's views on Russell's has caused irrelevant points of argument.

The argument I gave assumes that what Frege explicitly said should be taken at its face value. And why not? Indeed, the naturalness of the way Frege handles descriptions that express the senses of proper names is telling. Proper names involved in the descriptions are perfectly acceptable expressive elements of those senses they are part(s) of. I think that it is obvious that there is no slip on Frege's part. The explicit involvement of proper names in the descriptions of the senses of other proper names is much too striking for anyone to have failed to recognize it - let alone Frege with his logical acuity. It is much more likely that if Frege did not mean his explications of the senses of proper names to be taken as he gave them, he would have said so. (But, then, why should he have explicated them in a misleading way?)

The argumentative point of these examples is that the descriptions expressing the senses of proper names are not conceptually pure. It also puts in severe doubt the semantic usefulness of the general and/or qualitative representations of the referents. It seems to me that the doubt is so serious indeed that the motivating substance behind the Salmonian charge evaporates. Consequently no Fregean needs to honor the Salmonian demands on the senses.

Also Strawson, on the description theorists side, emphasized that the individuation and identification of the referent may involve non-descriptive and non-articulatable elements:

“[A speaker] may know the name of a thing and be able to recognize it when he encounters it, even if he can normally give no identifying description of it other than one which incorporates its own name.” (Strawson 1971a: 77)

“...an identifying description...may, of course, include demonstrative elements, i.e. it need not be framed in purely general terms. In general, indeed, it could not be so framed; it is impossible, in general, to free all identification of particulars from dependence upon demonstratively indicatable features of the situation of reference.” (Strawson 1959: 182, n. 1) (In fact the whole beginning of (Strawson 1959) could not be misunderstood: its emphasis is on referential communication and the role of demonstrative contacts with objects.)

The emphasis on the recognitional factor is what the neo-Fregeans have made quite much out of later. But independent of that I think we can conclude that it is beyond any reasonable doubt that the description theorists need not commit themselves to any “purely” descriptive and/or conceptual factors in identifying reference. Strawson calls such purely general descriptions “pure individuating descriptions” (Strawson 1959), like “the first dog born in sea”, but he denies that they would be very acceptable means, let alone adequate by themselves, for identifying references. In other words such descriptions are much too detached from what he calls “our general scheme of knowledge of particular things”. (That is, I think, precisely why such descriptions are so rarely used). And this scheme of knowledge is essentially spatiotemporal as comes to the identifying references made within its confines. Therefore the demonstrative factors are part of the descriptive means of reference. So contrary to what the direct-causal theorists claim there is a natural descendant line from Frege via Strawson to the neo-Fregeans, a line that emphasizes the partial i) non-conceptuality and ii) non-descriptuality of the senses of proper names. Consequently the demands for such purity by the direct-causal theorists are mistakes – indicating either poor homework or deliberate “straw-manning”.

The contextuality of the senses

The Burge-inspired Salmonian argument against the contextual involvement of the senses fails also. There is a compelling *Gedankenexperiment* by Frege that shows that the argument from contextuality does no harm to the Fregean theory. Frege gives the example in his letter to Philip Jordan (Frege 1997f) and it goes as follows. A geographer happens to see a mountain in one of his explorations. Upon that encounter he calculates its height and performs other such tasks the geographers are educated to

do. He then learns that the mountain is called "Afla". As it also happens another geographer is adventuring in the nearby area quite the same time and he performs similar geographical things but learns that the mountain is called "Ateb".²³ Frege then states the natural conclusion that the identity statement "Afla is Ateb" – formed when the two geographers happen to meet and exchange the results of their investigations - is non-trivial and informative because the two names differ in their respective senses. To support this it is in order to quote what Frege says on determining an object for it reveals clearly that the sense theoretic account involves the causal and the contextual factors.

"An object can be determined in different ways, and every one of these ways of determining it can give rise to a different name, and these different names then have different senses; for it is not self-evident that it is the same object which is being determined in different ways." (Frege 1997f: 321; emphasis mine)

It seems to me clear that both Frege's example and the quotation speak for themselves. Consequently the direct-causal theorists are very hard put to evade the conclusion that the explanatory machinery of Frege's theory, and by an obvious implication the Fregean approach in general, has always involved the causal and the contextual factors. As is easily seen especially the core elements of the baptisms and the naming situations are subsumed by the mountain example: there is no relevant difference between Frege's conception of determining the object and Kripke's notion of fixing the referent. It is true that Frege left open the way the relevant elements from the perceptual and the contextual encounters end up in the senses and the thoughts, and in what way they are effective in determining the referents when they are grasped. But that they *are* there can not be plausibly denied in light of the Ateb-Afla example.

Now we see that Frege's account of the senses is much more liberal than has been thought. The contextual factors are involved. The alleged gap between thought and world (through language) is in fact closed in Frege's theory. That some can grasp a sense, say, "born in Stagira", even if he has no causal or any other contextual means at his disposal to link him "directly" to Stagira, is not to be argumentatively jumped at. He has presumably learned from some source or another that Aristotle was born in Stagira, and as we can assume that the speaker is a competent language user, he knows or can infer that Stagira is a name for a place, a town most likely, and other such more general things in relation to the particular things about ancient Greece (and even about Aristotle). That is, he grasps the senses that constitute his grasp of the thought that Aristotle was born in Stagira. He may be unable to identify Stagira or find his way to Stagira without help, but that only shows that his sense of "Stagira" is quite limited. Those who live in Stagira or had been there have much richer senses of "Stagira". Most likely "Stagira" was introduced to the language of the Greeks (whichever its spelling was then) by someone who named it so; and some later users of the name had some other contextual means for getting to grasp its sense. But no matter what the details were, the Fregean theory is able to explain them.

It should be pointed out that the *neo*-Fregeans have taken it as an almost self-evident fact that the senses are informational and contain a fair amount of perceptual elements. But from my dialectical perspective this could not be just assumed, especially because I am arguing against the direct-causal theory. The reason is that it is part of the argumentative practice of the proponents of that theory to assume that the Fregean

senses do not contain contextual and perceptual elements. That is why they are thought to be "purely" conceptual and contain, or express, only qualitative/general properties. But the Ateb-Afla example shows that assumption to be an error.

Also on the description theoretic side the contextuality, hence non-pure-conceptuality of the descriptions (expressing the senses) was the official view before Kripke's arguments. Strawson maintained that in cases where the individual referred to could not be demonstratively identified in the context "...it may be identified by a description which relates it uniquely to another particular which can be demonstratively identified." (Strawson 1959: 21) (Context for Strawson includes "...at least, the time, the place, the situation, the identity of the speaker, the subjects which form the immediate focus of interest, and the personal histories of both the speaker and those he is addressing." (Strawson 1971b: 19)) It is easy to see that this demonstrative identification subsumes the naming situations (baptisms, dubbings, groundings). Consequently there is in this respect nothing new in the direct-causal theory's causal-communicative component, either.

With the contextuality of the Fregean senses that the Ateb-Afla example brings with it we can also reject another widespread misconception of what the Fregean sense theory is like (or, even, as it is sometimes implied, what it should be like). That conception concerns the "guiding direction" of the relation between the senses and their referents. The operative mechanism of the senses has been taken to be one-way street from the senses to the referents. That is, one as it were searches with a sense through individuals in a particular domain (in principle a whole possible world) and picks up the one who, or what, fits the sense uniquely. It need not be denied that there are obvious cases when one literally searches for the "fittest" individual with respect to the sense. (If that did not work, D.C.I.s would be largely unemployed.) But the point of the received conception about the Fregean view is that the process could not go conversely; it is not allowed that one acquires the sense from first encountering the individual that is, or becomes, the unique referent of the sense. But as we saw, this conception is misguided. The Ateb-Afla example makes it evident that there are no prohibitions for the speaker to have the individual at his disposal first and then for him to acquire the uniquely referring sense from that encounter. Afterwards the sense is used in identifying references to that very individual.

The neo-Fregeans have used the notion of the information about the referents of proper names. Being about, or of, the referent the information is supposed to derive from the referent causally. The only point I want to register is that the apologizing tone sometimes associated with this view by the neo-Fregeans in relation to the direct-causal theory can be forgotten. That is, the neo-Fregeans need not look at their view as involving a concession to the direct-causal approach with respect to the causal element. As we just saw that element originated already with Frege and therefore even preceded the direct-causal approach.

There is a double irony facing the direct-causal theorists. We saw that the causal and the contextual factors in learning a name, and of course in giving a name to something, belong to the theoretical machinery of the Fregean theory. But what I have argued reveals also that the communicative factors, like the reference borrowings emphasized by the direct-causal theorists, are nothing new to the sense theory. And generally speaking, it does not even seem plausible that the sense theorists would have thought

that the contextual, causal and communicative factors did not have any role in explaining reference, and referent retention - quite the contrary. The description theorists, or at least Strawson, and the neo-Fregeans have always taken those factors to be so obvious they have not emphasized them in particular. (But as I mentioned above Strawson was very explicit on those factors.) Rather the factors have become along with their arguments and examples tacitly understood to be there. (And tacitly thought that that is understood.) This is evident in, for instance, what Strawson says on the relations between the speaker and the hearer in those situations where the former is making an individuating reference (Strawson 1971: 78-81). I will not repeat his discussion here, I only point out that it does not need a trained eye to see that it is obvious that the communicative reference transmission is assumed there. For one thing how otherwise would it make sense to speak, like Strawson does, about the possible referential failures that take place in such communicative situations.

The direct-causal theorists deserve credit for their emphasis on the role of communication and the causal, albeit informational, factors in the referential use of proper names. But the use they have put them and the "picture" they have painted of both of them is much too arid one. This we have been seeing during this chapter time and again: the varieties of information related to proper names are overlooked by the direct-causal theorists. So I think that we can firmly conclude that the arguments from the demands of the senses being "thoroughly descriptonal" and "purely conceptual" and expressing only purely qualitative and/or general properties are based on a widespread misunderstanding of the Fregean theory. The Fregean theory *is* an internalist one in that it requires cognitive senses as the determiners of the referents. In this role the grasped senses form the epistemic boundaries of the knowledge of the communities of the speakers at any moment, as well as the knowledge possessed by every speaker individually. This determination of the referents is weak in that it does not automatically differentiate between the actual distinctness of the two (or more) referents in the Twin Earth scenarios. The reason is that in those cases there is not (yet) sufficient knowledge indicating the referential distinctness of the two terms in question in the two communities. But that does not amount to any kind of failure on the sense theory's part. The internalism of that theory *takes* the environments and the contexts of the thoughts and the referential expressions into account. The Fregean theory presents a midway solution between the extremities of externalism and internalism, respectively, i.e. that external objects are elements *in* the meanings or that external object has *nothing* to do with the meanings. The Fregean content is not externalist in the former sense. Before the discovery that there are two Huberts the contents of the thoughts of both the Earthians and the Twin Earthians were the same (by hypothesis). That is, what is wide and what is narrow depends on what is known at the moment the thought involving the referential representation is entertained. But neither is the Fregean content completely narrow for it takes the contextual factors into account in its very constitution. The Twin Earth arguments are based on too strong presuppositions of what the sense theory is committed to. Moreover the presuppositions of the arguments themselves are much too strong epistemically.

It seems to me that we can also conclude that there *is* a uniform, full-blown Fregean sense, fulfilling the functions of the subsenses simultaneously. Sense₁ as the conceptual content is also the way the referent is determined, the sense₂. The contextual and the causal elements belong to a sense of a proper name, either those acquired in the naming situations and groundings or those acquired later from other

sources, because these types of senses involve the cognitive relations to the contextual and object-dependent elements. (It follows also that the "purely" qualitative senses are not such: for instance "the tallest man alive" is contextual in relation to the humankind.) I have not dealt with the propositional attitude contexts here where sense₃ is taken to be prominent. But I do not see any difficulties such that Burge argues for. That two different thinkers do not believe the same thing when they believe that *p* because their respective senses of *p* differ from each other is not really a problem at all. Burge is misled (and many others, for this mistake is surprisingly common) in taking too seriously the notion of believing the same thing. To invoke again one of the main points of Frege's Aristotle note: as long as the intersubjective variance between the senses related to the object of a propositional attitude is not so wide as to prevent mutual comprehension it causes no troubles. In other words we should take the notion of believing the same thing (in general, "propositional attituding" the same thing) with a large grain of semantic salt. In fact so large grain that in typical cases it demands only, as Frege maintains, that the referents, or extensions, be the same.²⁴ So sense₃, the information value or the content of one's thought, does not differ from sense₁₊₂. From which it follows that the unified full-blown Fregean sense is a sound notion. Hence the whole line of the Burge-Salmon argument fails.

The imputation of the extreme version of internalism to Frege's view is understandable though. Frege was interested in the thoughts and the senses and held them to be "third realm" entities, distinct from the material objects and the subjective mental states. But in light of this it may be more adequate to say that the tendency is non-externalist rather than internalist. It is true that for Frege *Bedeutungen* are indispensable for the truth-values and truth-evaluability of the thoughts, but the actual procedures for finding out the truth-values has nothing, or not much of relevance, to do with Frege's logical and semantical framework. Senses determine referents, but it is the business of every particular scientific discipline to find out whether the thoughts about its domain of study are true or false. As already pointed out, the true thoughts are not the only thoughts, for it was Frege's contention that if they were, scientific research and questioning would be impossible.

The arguments that demand pure conceptuality of the senses are then based on misconceptions. In concluding this section I think it is in order to correct one more instance of this mistake. It is from an influential handbook on philosophy of language. Devitt and Kim Sterelny give a prompt formulation of the argument that the Fregean descriptive sense theory is inherently incapable of explaining reference, thus linking language to the world. They argue that the sense-cum-description theory only invokes other expressions, descriptions, but it does not give any account of how these are in their turn related to the world (Devitt & Sterelny 1987: 51-2). But the sense theory is fully capable of doing that as the discussion of the Afla-Ateb example above shows. Moreover this happens in the way that the direct-causal theorists cannot even deny because it involves the very causal and the contextual factors they emphasize. Devitt and Sterelny give a diagnosis from their argument. According to it the essential incompleteness of the Fregean-cum-description theory, in the form of the lack of the ultimate explanatory connection between language and world, was not noticed until recently – thanks to the direct-causal theorists. However, taken at its face value and in light of the argumentation above this diagnosis is quite incredible. It seems to me that instead Devitt and Sterelny, as well as the other direct-causal theorists, should rather have paused to reflect how unlikely it must be that such an easy point, the relation of

the descriptions to the world, would have been missed all along by the sense theorists. But as we saw above it has not been missed because it has been taken to be evident.

Internalism is best comprehended when it is realized that the representational capacities of the brain copy the external world. By this I do not mean the old and erroneous theory that mental representations are exact copies of the corresponding external entities. There are no copies in the brain in that photographic sense. But the brain does copy the external surroundings, from small to large, because of its informational connections from and to it; action is at least as important to representation as perception. The representations, the senses, the brain constructs are better understood to be structurally and dynamically isomorphous to the external entities. (That is the correspondence.) Of course this is a very general hypothesis, but which I am trying to begin to answer in chapters 7 and 8. As to its general, i.e. philosophical, substance I do not even pretend to claim any originality. But now its function is to help us in comprehending the naturalness of the internalist perspective. So if we could at least tentatively accept that the brain houses a copy of the external world – or rather many representational copies of the external entities like objects and their properties and relations, events and their interrelations; in short, the whole of the senses individually possessed – the externalist statements can be seen in the correct light. Take for instance the claim that the internalist narrow content is independent on the external surroundings and therefore inherently incapable of truth-conditional representation. But if it is understood that the informational constraints on the representations that the senses confer result in epistemic constraints also (=ESP), the position emerges in the midway between extreme version of both internalism and externalism. The mental representations *are* semantically related to the external world because they are informational. But that informational impact, with respect to individual representations, is not forceful by itself to support the externalist claims like “meanings ain’t in the head”. I argued above why that is not so: the content of an individual representation, the Fregean sense, depends on the other representations – which also contain information about their connections to the external entities they are about – and are therefore cognitive and therefore epistemic in turn. *Ergo*: reference is cognitive-cum-epistemic phenomenon. (Thus the content of the word, say, “water” is not, repeat is *not*, only H₂O and does not refer, repeat does *not* refer, rigidly to H₂O *before* it is known that water is H₂O.)

The content variance argument

According to Salmon’s argument from content variance the conceptual content is not part of the information value of a linguistic expression. The reason is that the content varies from person to person because people have different conceptions of the bearer of a particular proper name. These conceptions are also acquired in many different ways, from being acquainted with the referent to just fleetingly hearing something said about it. If the conceptual content is part of the informational value it follows that the informational value differs from person to person. But this clashes with the “original, natural idea” of a sentence encoding the same single piece of information for everyone (Salmon 1986: 68). Hence the relation of encoding information is objective and semantic, claims Salmon.

Now this argument could be attacked head-on by focusing on the semantic framework and the assumptions that Salmon uses. For one thing it is not at all clear that the notion

of information that best applies to semantics is the cognitively purified, not to say cleansed one, that Salmon uses; hence it is not clear at all that there is any “original” and “natural idea” of information encoding. Or that the only acceptable and semantically respectable entities are singulars (objects), properties and relations. But I will not adopt this strategy. (Though from my point of view Salmon begs some essential questions that have to be asked and answered before we pursue any semantics.) Instead I will argue that the first of the two caveats that Salmon makes not only robs the force of his argument from content variance, but also leads to an argument that reveals that one crucial question has indeed been begged by the direct-causal theory time and again.

As the reader remembers Salmon’s first caveat is that the encoding relation has to be relativized to a particular use of a sentence, of which the proper name in question is a part. (Kripke makes in effect the same remark (Kripke 1980: 8-9).) This relativization means in effect that the bearer of the proper name *must be individuated beforehand*. Take the sentence "Aristotle was fattish". It must be relativized, i.e. it must be decided whom the proper name is about. Until that is done no precise piece of information is encoded and expressed by the sentence. Thus if one is talking about the Greek shipping magnate, the sentence encodes the information that Aristotle Onassis was fattish, and the information value of "Aristotle" is Aristotle Onassis. Whereas if the sentence is used to talk about the ancient Greek philosopher it encodes the information that Aristotle the philosopher was fattish, and the information value of "Aristotle" is Aristotle, the commonly known Greek philosopher (though this last characterization is not part of the information value). That is what the content variance argument is all about; the conceptual associations expressed by the descriptions (like "the commonly known Greek philosopher") are irrelevant to the semantic attribution, to the encoding of information and having informational values – only the object and its properties matter.

Kripke says that the “classical description theorist” talk “for simplicity as if names had unique references” (Kripke 1980: 8-9). This is meant to imply that Kripke takes himself to be entitled to that simplifying assumption, too. But Kripke fails to mention that the “classical description theorist” (as any kind of description theorist) is entitled to that assumption precisely because it *follows from his theory*. The definite descriptions individuate and identify the unique referents (or at least they purport to do that in the actual practice of their use). It seems to me that this slip on Kripke’s part is indicative of the rather shallow general interest the direct-causal theorists have paid to the description theory (and to the Fregean theory). In fact I am even quite appalled here, for to me it seems obvious that when the “simplifying assumption” (the relativization of use) is made, it should have immediately been asked why is that assumption/relativization possible in the first place. Why have the direct-causal theorists turned the blind eye to the individuation issue involved? And especially after Kripke explicitly pointed out that only after the simplifying assumption can the rigidity claims be made at all.

Scott Soames discusses widely about the information expressed – singular propositions – and information conveyed, information depending on what the hearer knows about the referred things, by the declarative sentences/utterances (Soames 2002). But he premises his account on the “...simplifying assumption that ordinary proper names are unambiguous...” (Soames 2002: 55). However this again clearly amounts to bypassing

the more basic level of explanation that tries to account for the fact how one is able to single out a certain particular individual to be talked about in the first place. In other words, Soames' approach does not answer the problem of the relativization of the use of a proper name but presupposes that the name has been disambiguated already (in a context of use). Soames does explicate the competence conditions for the proper use of proper names by requiring that one either has acquired a referential intention that determines *o* as the referent (of "NN") - and this takes place by reference borrowing from the speakers (who intent to use "NN" to refer to *o*), or by acquaintance with *o* (Soames 2002: 65). But as I already argued the first subcondition is much too weak, for intentions to co-refer do not guarantee the retaining of the referent. And the acquaintance with the referent is allowed by both the Fregean and the description theories, as we also saw above.

I endorse the caveat myself, but I argue now that the caveat in fact disarms the content variance argument. The point of my objection is that if the use of a particular proper name has to be relativized to a particular individual out of the many others with the same (syntactically individuated) name, that very individual must be identified in the relativization process. Put in another way individuating the bearer of the name is in effect already making an *identifying reference* to that individual. Still in other words, the argument from content variance *presupposes*, and even in explicit terms, that we must first be able to identify the individual for the Salmonian semantic framework to be applicable at all. But that is a cognitive task, hence readily amenable to the Fregean approach.

The direct-causal theorists are likely to argue that the relativization of the use, i.e. the individuation of the referent, is taken care by the groundings and the causal networks. But that argument fails by now. I have spent the larger part of this chapter in arguing that the causal factors by themselves are insufficient for any referential purposes. It is the information they transmit that is the *primus motor* in reference. Neither are the intentions to co-refer explanatorily sufficient, for they depend on the information the speakers have access to, even if eventually via deference to the information the experts possess. Furthermore Salmon himself rejects the causal connections as individuating factors as they are postulated in Devitt's theory. He takes that proposal to be a category mistake (Salmon 1986: 70-1). So I conclude that the referent could be individuated only by invoking information about it; that is, by invoking the Fregean senses related to the proper name in question.

Because this is the main point of my argument against the sufficiency of the communicative links *qua* causal it is in order to bolster it with an example. The direct-causal theory claims that no (non-trivial) pieces of knowledge, or beliefs, are necessary elements in successful reference; instead the communicative causal links suffice for reference and irrespective of whether we know their previous courses or not. (That is because we very seldom do.) The communicative aspect is emphasized because the direct-causal theorists claim that their theory explains the ways we commonly express thoughts and use proper names to refer. So they will be under severe theoretical pressure if the communicative facts do not confirm the theory on that count. Now witness the following rendition of an exchange that took place a while ago.

-“People should not want too much. But everyone should have so much that he feels good.”

-“That is pretty much what Aristotle said.”

-“Who is Aristotle?”

Then I told the wise 4 years old who Aristotle was: that he was an ancient Greek philosopher; that he lived 384-322 before year 0 (she likes to say “year zero”); that he thought that moderation in everything one does is the best way to live; that he is taken to be the most brilliant of the Greek philosophers; that he was a pupil of Plato (but about Plato we maybe talk some other time). I am not sure that all of this registered, especially as her conception of the ancient Greece might be somewhat hazy. (Mine is, anyway.)

Now the direct-causal theorists may claim that the child’s use of “Aristotle” was relativized nevertheless because she automatically deferred through me. I do not wish to contest that that is what she can be said to have achieved. But as I already argued deference does not complete reference make. In particular she did not make genuine speaker reference to Aristotle at the moment she asked who Aristotle was. (That only seems so if one could not shrug off the highly idealized and omniscient perspective the direct-causal theory works with.) Her question did manifest the particular tacit application of the metasupposition, i.e. that “Aristotle” is a name of someone. But this metasupposition does not by itself tell anything substantial about the referent. What is left then? What is left is to understand why she asked who Aristotle was. The answer is obvious: she did not know at all who Aristotle was; nay, she did not have even rudimentary information about Aristotle. So there was nothing with which she could have individuated the referent of “Aristotle”. In other words her use of “Aristotle” was not yet “relativized” to anything. And if the direct-causal theorists keep on claiming that “of course the use was relativized” because of the deference through me, the answer is that *I* have had to individuate Aristotle the philosopher as the referent of “Aristotle”. Either way the need for the individuating information about Aristotle the philosopher will not go away.

We see here again how the direct-causal theory meshes the epistemic perspective of those “in the know” about the referent with the highly idealized picture of the causal routes. In fact that is doubly erroneous: in addition to the unrealistically cleansed causal picture (and *its* tacit mesh with the elements of the demonstrative references due to the basic schema of reference in action) the direct-causal view invokes knowledge, but which according to the official doctrine should play no role in the reference of proper names. It may be in many cases, or even in most, that in addition to the activation of the basic schema of reference, the tacit involvement of the metasupposition and also the mere knowledge of the existence of the practice of deference all contribute to the so far unscrutinized “intuition” that is claimed to give evidential support to the causal explanation of reference. The latter two explain in turn, but only in part, the prevalence the perspective of the semantic reference is usually given over the speaker reference. The other part consists of information: the shared knowledge about the referent - together with the shared knowledge of that knowledge being shared. But as we saw in the discussion of the Feynman case it does *not* follow that when there is semantic reference with a particular proper name the speaker reference conforms to it automatically. To claim this as Kripke (with his direct-causal followers) does, is to confuse the above factors with the role the individuating information about the referent has in the speaker reference. Therefore a proper name does not speaker refer if the speaker does not know sufficiently individuating information about the referent.

I think that my argument gets even more support when the perspective of the ignorant speaker is adopted. Many of us have been in a similar situation: first completely at a loss to what the speaker is referring by a certain name. Usually the topic of the conversation affords some clues where to “locate” the referent spatiotemporally. But we may imagine that the hearer just arrives at the scene and only overhears the proper name used. When we put aside the application of the metasupposition, as I argued above that we should, we see that the hearer is in a hopeless position referentially speaking as long as he does not acquire information that would help him to individuate the referent. (Recall the Liebknecht example in note 11.) In short, I do not see how we could escape drawing the conclusion that in the ordinary referential uses of proper names it is the information related that only makes it possible to refer genuinely.

We saw above in section 4.3.1 that senses do determine referents, but this does not involve any stronger mode of determination than what the knowledge and beliefs possessed by the language users allow. If the epistemic state of the community changes, i.e. if the knowledge states of some, or even one, of its members change upon the discovery that what they thought about, say, Hubert is not in fact so, that will have its effects on their referential practices – though not as immediately as is thought by Putnam. I would not argue the next point in detail for it would digress too much in this stage, but it seems to me that Salmon incorrectly distinguishes the information encoded by an assertive sentence as if it were independent from what the speakers know who assert the sentence, or who use the sentence to refer to some definite individual. "Aristotle was fattish" certainly encodes information that can be expressed by saying something like that the Greek philosopher commonly known and called (in contemporary English) "Aristotle" was fattish. Put in Salmon's (Russell-Kaplanian) jargon the singular proposition expressed by the sentence consists only of Aristotle himself and the property of *being fattish*. The mode of presentation of the referent is *not* part of that proposition. This type of proposition is all right in semantic purposes at least in that it seems not to be an incoherent notion. The notion even seems to have a role to play in explaining out truth-seeking activities. But it has to be noted that the singular propositions are not what primarily count in this context. What does count is the way we *get* to the propositions, the way their constituents are, as it is said, loaded into them, and that they indeed are the intended constituents. But this *requires* the identification of the individuals (and individuation with respect to properties and relations). Singular propositions and their relations to truth-conditions of the sentences, which express them, owe what explanatory fruitfulness they may have to the epistemic and the cognitive factors that are invoked in constructing the very propositions to begin with. Thus it may be that the singular proposition that the descriptive sentence like "the second king of France was fattish" expresses consists of a certain individual and the property of his being fattish, not the Fregean mode of presentation "the second king of France" and the property of being fattish. But who that very individual is (*Charlemagne*) must be predetermined by some epistemic and cognitive means for the singular proposition to be of any use in communication (and in truth-seeking practices). And these means are the Fregean senses (which the descriptive modes express).

It certainly seems strange to hold that a flesh-and-blood individual is *part* of a proposition. Frege never accepted that. In many occasions (for example in his letters to Russell) he complained that some (like Russell) think that, say, the Mont Blanc itself is a part of the thoughts about Mont Blanc. The above elaboration of my argument about

Salmon's caveat should be taken as providing the justification for Frege's view. We see again how the Fregean and the direct-causal theories differ from each other. Singular propositions are the pet entities of most of the direct-causal theorists. The reason is that the modes of presentation of the referents are the conceptual counterparts of the descriptions, and since the direct-causal theorists have disregarded the semantic relevance of the descriptions for the reference of proper names, the singular propositions are the natural choice for them because the modes of presentation are not parts of such propositions. But this ontological move only betrays the cognitive poverty of the direct-causal theory once again. If there has to be - as there has - the relativization of the use of sentences to be about specific individuals at a time, that re-enters the Fregean factors with full force. Also on this score the *horror cognoscere* haunts the direct-causal theory of reference to its peril: it could not provide us with genuine explanation of the workings even of its own basic postulations.

Another way the direct-causal theorists could still try to counter my arguments is by stating that their examples and arguments are used to help to explicate the conceptual aspects of our notion of the reference of proper names. Therefore it is sound methodology to impose our epistemic state and perspective on the examples. The examples also clearly indicate that reference is causal. Unfortunately I could not accept this view either. For the conceptual aspect not merely to coincide with causal "intuitions" something must differentiate it from the latter to give it even a minimal credibility. But if my arguments here and in section 4.2.5 are to the point there just is no explanatory substance left with the "conceptual" aspect of the notion of reference relating to the causal links. When these arguments are coupled to the other arguments showing the overall explanatory poverty of the causal links (in section 4.2.3), the direct-causal theory is seen to be explanatory failure.

The direct-causal theorists can of course retreat and claim instead that the conceptual aspects reveal the (what I called) formal, or pragmatival, aspects of the reference of proper names. But because the Fregean as well as the description theory can incorporate those aspects - because of their very "formality" and pragmatic status - the conclusion is that the direct-causal theory is only a subtheory of the Fregean theory of reference. We will also see a little later that there is another line of argument that shows how this result is the best the direct-causal theory could hope for. So all in all we can conclude that the Kripke-Salmon-Soames caveat line pulls out the argumentative rug under them: if the individual the proper name in question refers to must be singled out "in the context" before any semantic results could be achieved, in particular such as are used in arguments for the rigidity of proper names, their game is lost to the Fregeans. That is, the bearer of the name, in the context, has to be individuated and identified *before* any direct-causal considerations could even to be applied to that name. Since it could not be the causal links that enable us to do that, what is there left but the Fregean senses-cum-information?

Rigid designation and naming

From the relativization argument follows a corollary that is of utmost importance to our subject. It is that Kripke's idea of proper names being rigid designators also presupposes the individuation of their referents. A proper name is a rigid designator if it refers to the same individual in all the possible worlds, or counterfactual scenarios, in which that individual exists and that individual is the actual-world referent of the

proper name in question. According to Kripke this explicates the modal difference between proper names and the definite descriptions, for the latter refer to whatever has the properties the description specifies in a particular world under consideration. This is a stipulative move on Kripke's part (as he acknowledges himself). We start with the actual referents of proper names in our language, not with the possible worlds. In other words we do not, as it were, go to a possible world and try to find out who or what the referent is satisfying the properties the description gives. We have the referent firmly fixed to begin with. This stipulation is an attempt to pre-empt any counters that use the notion of transworld identity. The application of this notion requires that we can make sense of the transworld identity of individuals antecedently in our counterfactual talk. But to Kripke this is to put the cart before the horse: "it is *because* we can refer (rigidly) to Nixon, and stipulate that we are speaking of what might have happened to *him* (under certain circumstances), that 'transworld identifications' are unproblematical in such cases."²⁵ (Kripke 1972: 270) But one could not fail to notice that this is just to trade the problem of the transworld identity for the problem of actual identity. Salmon makes the same trade for he says that "in the case of nondescriptive singular terms, there is no such possible world relative process for determining denotation."²⁶ (Salmon 1982: 33, n.35) But that is patently false: actual world is also a possible world; so it is the actual individual we are referring to that must have been antecedently individuated (or at least during the process of talking about it). It seems that in effect Salmon also says that the mechanism by which the free variables are assigned values is a paradigm case of the non-Fregean semantic mechanism. But this only hides from the view the fact that the assignment presupposes an *evaluation function* by which the values are assigned; and that function is all that a Fregean needs for his purposes. In fact such evaluation function may be taken to consist of whatever the Fregeans need for their explanatory purposes (see the second quotation from Almog below).

As we saw in the preceding section virtually in all cases when the direct-causal theorists discuss the reference of proper names their characterizations are like this: "x" refers to *x*, and it is then made clear, or let be understood from the context, that *x* is some well known individual, i.e. already individuated. It is easy to see that this is in effect Salmon's relativization caveat again. Joseph Almog has observed a related point about the antecedent individuation, for he says that

"On the possible-worlds picture, a rigid designator, weak or strong, *de jure* or *de facto*, is defined as a designator picking out the same individual in all possible worlds (with this or that qualification). But for this to get off the ground, the idea that the transworld identity of individuals is given must be taken for granted. Question it, make the very idea of overlapping domains meaningless...and rigidity is tottering." (Almog 1986: 227)

It is in order to note, however, that Almog's own view, his "structural objectivism", presupposes the "generation" of propositions (in fact singular propositions with the "flesh-and-blood individuals" in them (Almog 1986: 220, 225)), and therefore could not tell anything about the individuation of the referents of proper names. As I argued the antecedent individuation is required for us to be able to generate the right singular propositions expressed by the sentences. In other words one has to know who, say, Quine is in question when one generates the proposition the sentence "Quine is human" expresses. Of descriptions Almog realizes that they are not naming devices but devices that induce conceptual contents at the generation stage, "blueprints" for the

referents (Almog 1986: 225). The distinction between the generation and the evaluation of the propositions as to their truth-values in possible worlds seems to me obvious in this type of semantics. Almog presents this (what seems to me to be) under-the-carpet-individuation with respect to proper names in a somewhat different way. He speaks about Kripkean "one-stage theories of evaluation" and says that

"a point of evaluation...may include a world, a time, an agent, a spatial location, or what have you. *No matter what information is needed for evaluation*, it can be packed into a single 'index' (with many 'coordinates'). Then, setting a sentence relative to that index, we obtain the sentence's truth value." (Almog 1986: 218; emphasis mine)

So although Almog perceives that without the individuation of the referent we could not achieve an adequate semantics, he fails to see that the point of the evaluation lies just where the Fregean factors are to be found. In other words, not even Almog does realize that the direct-causal theory presupposes the Fregean theory.²⁷

Almog *does* say that the general Kripkean one-stage evaluation framework *can* provide for the semantic notions like contents, so presumably also the senses. As he points out the content of a sentence is its evaluation function that assigns to it truth-values in accordance with how things are in those possible worlds with respect to which it is evaluated (Almog 1986: 214). But as the formulation itself shows this is in effect just what the Fregean senses of the referring expressions do: they lead us to the referents, to the things in the world(s), and so enable us to judge the truth-values of the statements of which the expressions are parts.

As we saw earlier Soames admits that there are partially descriptive names (Soames 2002). For instance "Mountain McKinley" and "President Halonen" are such names. (As these two examples show the Soamesian partially descriptive names should be kept separate from the Evansian descriptive names like "Julius" which is the name for the one who invented the zip, because the latter names have their descriptive content related to them implicitly at the linguistic level.) But it should be obvious that the partially descriptive names indicate, in being halfway between proper names and the descriptions, that proper names *do* have senses expressible by the descriptions associated to them. The partially descriptive names, so to speak, have pulled up some descriptions from among the rest of the class they belong to, with respect to the proper name in question, and characterizing the whole sense of a name. That there are those other descriptions, i.e. that the particular sense is informationally richer, is shown by the fact that for one to understand whom I am talking about when I say, for example, "president Halonen gave a speech today", the hearer needs to know also which country Halonen is the president of and other such related things. That proper names are not semantically different from the partially descriptive names should be evident because one may say that "Halonen gave a speech today" and some other one would understand what was said because that someone comprehended that most likely one was not talking about his eccentric neighbour, but the president. The descriptive part of the proper name is then seen to be optional and its use context dependent.

The category of the partially descriptive names, as Soames conceives it, is rather weak one in that such expressions are rather seldom used (except when the referent has to be disambiguated, or it is the practice to mention the descriptive element, as with the

presidents). Those names are quite poor in their descriptive content because the predicates do not by themselves contribute the unique identification of the referents by themselves in every case. Take for example the descriptive name "Professor Saul Kripke" (one of Soames' examples). It does not pick up any professor SK in case there happens to be more than one such. (Or it picks them up distributively.) I think that this does not merely indicate but in fact shows that the sense of a proper name, as uniquely referent determining, is richer. The reason is that Soames' very examples would not work if they did not indicate other, tacitly presupposed, parts of the senses; for example that it is professor SK, the famous philosopher of language and modal logician, the author of *Naming and Necessity*, and such other pieces of descriptive information.

Now that we have swept the causal factors aside together with the "intuitions" there are left only the sense theoretic factors for the individuation of the referents of proper names. We can accept that proper names *are* rigid designators, but this just shows that the notion of rigid designation could not be used in the arguments against the sense theory of reference. It follows that the idea of rigid designation must presuppose the sense theoretic individuation of the referent for its meaningfulness. Put in other words the direct-causal theory only seems to be a viable theory of the reference of proper names because it works implicitly with what only the Fregean sense theory can provide. The result is that rigid designation belongs at the "phenomenological" level of reference of proper names. But this does not abandon the sustainable idea that we can "actualize" the descriptions, i.e. fix their referents by an "actuality operator" so that they pick up the same actual referents proper names do in all the relevant possible worlds. This way the descriptions become as rigid as proper names. But rigid designation does not reach into the core, to the explanatory level any theory of the reference of proper names has to provide. It is precisely at that level where the identification and the individuation of the referents are operative. By the same token it follows that we can accept that proper names are directly referential expressions, but they are such only because the Fregean mechanisms of the individuation and identification of the referents have had their say in the "relativization of use". But proper names are *not* directly referential if by that is meant that the referents are not determined by the Fregean senses. Proper names are directly referential only in that they behave differently from the non-actualized descriptions expressing *parts* of the senses of proper names.

We saw in section 3.4 that Recanati draws a distinction between rigid designation and pure referentiality (or type referentiality). Before him A. Smith had discussed the distinction (Smith 1984). Because his discussion helps to develop my view further with respect to the direct-causal understanding of proper names and their primary mode of functioning I will now shortly discuss Smith's paper. Smith argues for the semantic difference of proper names (without modal considerations) from actualized descriptions (to which we return in the next chapter). His argument is that it is the notion of *pure referentiality*, not of rigid designation, that brings out the difference. In essence the rigidified descriptions denote their referents by satisfaction, i.e. by the referents being the objects the world contributes to the denotation relation. But proper names denote without any satisfaction relation; a proper name designates, *names*, a particular individual. According to Smith with names

"the relation of term to item in the world is internal to semantics...because names...do not first express a descriptive condition...Since the fact that a certain name names a certain individual is internal to the meaning of that name..." (Smith 1984: 190-1).

This is the crux of Smith's notion of pure referentiality (of proper names). But this view does not succeed. We can take it for granted that the semantic descriptive condition the description expresses has to be met for that description to refer to an individual. But, as Smith says, proper names have semantic meaning condition attached to them also, and that condition characterizes their referentiality: it says in effect that they are names of individuals. Hence there is no *functional* difference between the descriptions and proper names from the referential point of view. The descriptions refer by their descriptive condition, proper names refer by the condition that they are names. As I have kept pointing out the crux of the matter is that we still has to know to whom a particular proper name "NN" refers, for there may be many other NNs around (the problem of the disambiguation of proper names). And the proper name by itself does not tell us that. So in addition to its naming condition there has to be related some other condition which enables us to refer to a certain NN and not to any other NN. It does not seem to me to be of semantic and referential relevance that proper names are, so to speak, one remove from the descriptions. In other words, behind any successfully used proper name there stands a description (or descriptions) expressing the sense of the proper name.

No one denies that proper names name, and no one denies that the definite descriptions describe. That proper names behave differently from the mere descriptions, i.e. the non-actualized descriptions, in modal contexts (and in counterfactual talk in general) is true. That proper names are rigid designators, compared to the mere descriptions, is the explanation of that differential behavior. So, despite the air of triviality of the difference, I am not saying that Kripke does not deserve the credit for pointing out the modal differences. But I have no doubts whatsoever that if Frege and Russell had been interested in the modal issues in relation to thoughts and language, they would have spotted the difference readily. (It seems to me that Russell's distinction between the primary and the secondary occurrence of the description is the starting place for that.) Still we must be careful not to lose sight of the fact that the function of proper names is to name, both in form and in substance. That is, they name in form because their function is to pick up individuals. That is what their (pure) referentiality amounts to, or equivalently their possessing of Recanati's postulation "REF". But they are, or rather we as speakers are able to do that, to name in substance, *to refer*, because of the role of the "blueprints", the related senses (which the descriptions are used to express). Proper names are rigid because the named individuals must have already been individuated and identified for the names to be meaningfully and not just formally used (i.e. only as names). And this includes the counterfactual talk as well. Pure referentiality pinpoints the formal aspect of proper names, rigid designation the substantial aspect (ditto for the actualized descriptions).

In short naming should not be confused with genuine referring. A proper name names a certain individual, but it is able to do that (or strictly speaking the speaker is able to do that), to name *that* particular individual, i.e. to *refer* to him, her or it, only because of the sense of that particular name. The basic mistake of the direct-causal theorists as Millians is to equate naming with reference. A revealing case is Kripke himself.

"According to Mill, a proper name is, so to speak *simply* a name, it *simply* refers to its bearer, and has no other linguistic function. In particular, unlike a definite description, a name does not describe its bearer as possessing any special identifying properties." (Kripke 1979: 239-40)

Everything in this quotation can be accepted, but only if the confusion of name's simple function to name and its reference, i.e. how it is able to name a particular individual, is avoided. A name does not describe but names; if it described, it would be a description and not a name. But the name *refers* because it has descriptions associated to it expressing its sense. Everything is fine when we keep separate the two semantic elements related to proper names.

Conclusions

It is time to review the results reached in this chapter. The modal and the epistemological arguments against the Fregean sense theory of reference (or the "Fregean description theory") are misguided because the former presuppose 1) strong semantic relations between proper names and the descriptions expressing the senses to which neither Frege nor the Fregeans are committed; and the latter presuppose 2) such *a priori* knowledge of the referents of proper names which Frege did not and the Fregeans need not have anything to do with either. Large part of the negative attitude to the sense theory of reference is explained by these misunderstandings of what that theory is or should be like. The Fregean theory is an *empirical* theory of reference, not a philosophical "meaning" theory or theory providing analyses or definitions of proper names. What has happened is that these cherished methodological-cum-semantic assumptions from the 1950s and the 1960s have been forgotten to have been operative but not dropped, and what has been left is the too easily accepted impact of the first (Kripkean) arguments because of the forgetfulness. This is rather curious because it was shown fairly soon after the publication of Kripke's arguments that 1) the "actualized descriptions" view survives the modal argument, and 2) the causal component of Kripke's theory has problems of its own ("cut" and changed causal links). The other part of the explanation of the dominance of the direct-causal theory is the conception that neither the Fregean sense theory nor the description theory has anything to do with the contextual and the causal factors essential in reference of proper names. But as we saw the Fregean theory *does*. And so does the description theory: Strawson has discussed the preconditions of the individuation of individuals and these contain the contextual and causal elements (Strawson 1959).

The semantical argument fails also. There are both internal and general reasons for the failure. The internal reasons have to do with the explanatory poverty of the causal links, even when they are coupled to the communicative intentions to co-refer and to the (perceptual) groundings of the referents.²⁸ What explains referent retention is the information that the referential communicative exchanges transmit – i.e. to all intents and purposes the Fregean sense theoretic factors in modern terms. The general reasons have to do with the misunderstanding of the referential intuitions, on which the direct-causal theorists have relied rather heavily.²⁹ The intuitions telling that proper names have referential power of their own are in fact 1) only indications of the basic schema of reference at work, the cognitive and neurolinguistic functional structure on which

all referential use of language is based; and 2) the mistaken belief that proper names as rigid designators have no semantic connection with the definite descriptions. The direct-causal theorists have also a highly idealized picture of the very causal connections from the referents to all subsequent uses of their names. The only situations where the causal links are in force in the way that their robustness can be relied on are the demonstrative references. But the Fregean senses take these situations into account. The Ateb-Afla example demonstrates that the senses contain the contextual and causal elements and are therefore not "purely" conceptual. That this has not been realized explains, in addition, why the Fregean sense theory of reference ("Fregean description theory") has been deemed refuted. But here I would go even so far as charging the direct-causal theorists as having not been up to the standards of their trade: Frege's Ateb-Afla example has been available for quite a long time in Frege's published correspondence.

The causal picture is a very high-level picture and very idealized at that, too. In contrast the Fregean approach focuses on the "microlevel" of referential language use and the factors operative at that level: varieties of information gathered and stored in memories, books and other records. So we see that there is a serious rift. Because the theoretician's perspective of the direct-causal theorists is such high-level and idealized one, they are prone to claim that reference is independent on what the speaker knows or does not know about the referent; only the causal links matter. But the actual practice of referential language use is not nearly that clean; it is epistemically and informationally dirty, and it is local. The localness underwrites the informativity of reference. Reference is informational business; the talk about causal links misconstrues this from the very start. Consequently it is easy to succumb to believing that the causal links exist as such. As I argued they do not, and from this it follows that the transmission of information is what counts for referential purposes. Therefore we use expressions to refer because of what we know and believe – because of the information we possess (purportedly) – about the referents. In the case of mixed links information rules; Madagascar case shows that. In the case of unbeknownst cut links we can accept that reference is not always made, but even that depends on the amount and epistemic quality of the information about the referent "beyond the cut". The contrary "intuitions" are just borne by our demonstrative *Dasein* erroneously writ large.

To continue the diagnosis of the results of the failure of the semantical argument, apart from the demonstrative situations all other temporally extended causal connections are shown to be incapable of explaining the referential uses of proper names. The causal picture has been unclear because the first causal proposals (Kripke) mentioned communicative events as the nodes in referent retention but left the matter at that. In other words, the causal perspective was so general that it forcefully ignored the specific factors in those communicative events that really contribute to the retention of the referents. Also the talk of the intentions to co-refer and reference borrowings was left at too general level. In effect it masked the actual factors operative in referential communication. Therefore the direct-causal lore was able to survive and be taken as superior to the Fregean theory. (However some direct-causal theorists have realized that the informational factors must be given an explanatory role. But to my mind this amounts to inadvertently admitting that there is not much left with the original content and aims of that theory.)

I argued that the common but mostly tacit knowledge of the practice of deference explains many of the claims made in the counterarguments to the Fregean theory. I also proposed that this factor – together with the basic schema of reference as the precondition for any use of language – constitutes a general mechanism prevailing in the linguistic communities. If that is correct we can conclude that the direct-causal theory has, so to say despite of its proponents’ aims and explicit claims, been concerned about these *preconditions* instead of the *substantial* matters at the level of the actual practice of reference. It does not help to claim that that is what the reference of proper names is, then. For as I argued the Fregean theory involves the preconditions also. And the claim would not be wholly true, either: the Fregeans can explain the substantial issues that are also commonly taken to be part of the phenomenon of reference (like Frege’s puzzle, the propositional attitude contexts).

So all in all, the “Fregean description theory” has been an easy target only because the direct-causal theorists have forced it into a straitjacket of synonymies, *a priori* knowledge of the identities, purely qualitative/general properties and predicates, non-deferentiality and isolated from typical referential communicative practices (because its being allegedly unreasonably internalist). Not all of these presuppositions have been deliberate: as I explained, the theoretical climate in the analytical philosophy of language back in the 1960s contributed a lot. But the overall result has been that the direct-causal view has been able to grow on seeming importance at the expense of misconceived Fregean (and description) theory.

With the anti-Fregean and anti-descriptivist arguments it has become common to adopt the stance that we have to make a distinction between *purely referential semantics* and *cognitive semantics*; distinction between how we refer and how we think the referents (Wettstein 2004, Almog 2005). The issues here are more complex than that formulation reveals (for example there are singular propositions involved, empty names, and all usual problem cases one can find). But the general rationale for this distinction is that in the ordinary referential uses of language the directly referential expressions – proper names being the primary class – function differently from the way the Fregeans and the description theorists have thought. It is the received view that there are not Fregean senses to mediate and secure the reference, neither uniform descriptions related to the referents. Consequently it is believed that the only robust things are the causal-cum-communicative links from the naming events onward.

The cognitive semantics deals with the thoughts of the speakers and focus on such phenomena as Frege’s data. The purely referential semantics is, however, premised on the soundness of the anti-Fregean arguments. My general argument in the first part of this essay has been that this decoupling of the purely referential from the cognitive thought contents does not and *could not* succeed. As we saw the anti-Fregean arguments and the arguments allegedly giving support to the direct-causal theory fail one way or another. Hence the purely referential semantics is not an option. (Of course in the “logical space” there is a place for such referential semantics that is not an antagonist to the content semantics; i.e. purely referential semantics could be independent from the anti-Fregean reasons. But the corollary from my argumentation is that this option is empirically ruled out also.) The basic reason, or factor, that underlies this failure is information. This amounts also to the “could not” part, for I am arguing that referring as a form of human action is intentional, and because of that it is informational and cognitive. Information finds its way into proper names and it needs

to be individuating for genuine reference to occur. Towards the end of this essay I will develop the cognitive, and neural, approach to modernize the Fregean senses. But for now it suffices to bring out the general diagnosis of the basic error in the direct-causal-inspired referential semantics. It is that it takes reference and thought to be independent on each other. But they are not. There is a constitutive dependency from cognitive to referential.³⁰

It should be mentioned that the very general theoretical switch from the dependence of language on thought to the directly opposite way (or alternatively their mutual independence) is a possibility, at least as far as I can see. But what I deny is that the particular switch proposed for example by Wettstein and Almog could not be achieved for the reason that I have been arguing: the general theory of reference on which they rely is an explanatory failure (Wettstein 1991, 2004; Almog 2005). Of course it would be very satisfying if I could construct a general argument showing that no dependence of thought from language is nomologically possible. I do think that I have the makings of such an argument in what I argue in the appendix to chapter 6. It is an argument for the wholesale Fregean view according to which language expresses thoughts and these involve the “cognitive fix” on the referents - and that could not be otherwise because of our biological nature.

But let us go back to the main course of argument. In a way the direct-causal theory is a *retreat* from the sense theory. The direct-causal theory looks at the communicative links and nodes from a far too general perspective for it to be of help in the specific explanations of the referent retention. That is reflected especially in the reliance on the co-referential intentions without realizing that it is the information about the referents on which the intentions must depend for them to be able to fulfill their role in the reference retention. And because the Fregean sense theory has all along contained the communicative, hence causal, links as information transmitting, the direct-causal theory is no advance. The issue has been over what referential factors one focuses. The direct-causal theorists have focused on the communicative links and nodes, albeit – as we have seen – poorly scrutinized. The sense theorists have focused on the information about the referents, for that is what the talk of the senses is in modern terms. But all in all since the arguments for the direct-causal theory fail in the crucial places, as well as do the counterarguments to the sense theory, it can be concluded with sufficient firmness that the direct-causal theory of reference is explanatorily hopeless with respect to those factors that are not already parts of the Fregean sense theory.

The general result is, then, that the direct-causal theory of the reference of proper names, in regard of its few viable features like demonstrative references, is at most a subtheory of the Fregean theory of reference. This conclusion is based on our findings that i) the Fregean senses take into account the very causal and contextual elements the direct-causal theorists have thought are what set apart their theory explanatorily from the Fregean theory; and ii) the semantical viability of the idea of rigid designation presupposes the Fregean individuating machinery. Neither are the Fregean senses purely conceptual (and general, or qualitative) modes of presentation of the referents, nor need they be; they can be and usually are object-dependent and involve contextual information. That the much used Twin Earth arguments make epistemic assumptions which the Fregean sense theory need not respect strengthens the case for that theory more still. That is not to be wondered at, really, for the contents of our cognitive and epistemic states are so to say “local” at most and the “grasping” of the Fregean senses

happens only locally. That is why the reference of proper names is, *pace* the direct-causal theorists, essentially cognitive and epistemic phenomenon. And that is how it should be because referring is a form of human action, and all human action is intentional, i.e. cognitive. So we see that the direct-causal theorists have all along been Fregeans by another name. (Their mistake is that they believe that name is everything.)

The rather irritating implication from the direct-causal theory is just its causal robustness. This does not fit well with the facts. After the naming event – though already to some extent in them: ostension is not unproblematic pointing device, neither is the use of demonstratives - the fate of the causal links is at the mercy of the misunderstandings, forgetfulness and such vagaries on the speakers' part - just like anything we do. There just seems not to be available any stronger, or more basic means than the cognitive and epistemic ones to guarantee that the original referent is preserved as the correct and intended "source". But the direct-causal theorists seem not to care because they confuse 1) the information transmitted with the causal links as such, 2) the tacitly applied deferential metasupposition, together with the misgeneralization from the demonstrative situations of reference and perception of objects, overrules the speaker reference by the semantical reference. But that begs the issue in many cases.

According to the Fregean theory there may be much less reference around. That prediction follows from the nature of genuine reference being based on the individuating information about the referents, the senses of proper names. When the speaker does not have sufficiently individuating information at his disposal, the feeling that a genuine reference is made with the proper name is explained by the tacit working of the deferential metasupposition together with the activation of the basic schema of reference. (And that explains also why proper names name and why the definite descriptions can be used to refer.) But I have been arguing that this is *not* genuine reference, substantial reference, but only reflects the preconditions of the referential use of language. Genuine reference depends on the cognitive and epistemic factors and these depend on the sufficiently individuating informational factors.

Appendix: Devitt on the "four puzzles"

I have been arguing that all of the counterarguments to the Fregean sense theory of reference fail for some crucial reason or another. Now I want to change the gear diametrically and charge the direct-causal theory of reference with respect to its constant failure to explain the famous four puzzles of reference and cognitive significance. In other words the charge is about the *empirical* poverty of that theory. The four puzzles are the informativity of the identity statements, the positive and the negative existential statements and the propositional attitude contexts. (Acute observations about the issues can be found in (Wettstein 1986), (Bach 1987: 162-7), (Devitt 1989), (Salmon 1989) and (Nelson 1992: 154).)

Remember that the puzzles are motivated by the problems with the Millian and the neo-Russellian views according to which the semantic value (or the sole meaning) of a proper name is its referent. The informativity (the non-triviality) of the identity statements like "a is b" is a semantical problem for this view because, though "a" and "b" have the same referent by hypothesis, the statement tells us more than the tautologous "a is a" (or "b is b"): the cognitive significances of the two statements

differ from each other. Russell argued that the positive existential statements like "a exists" become trivial or redundant because "a" already names an entity, so there is really no need to affirm its existence (affirming becoming semantically idle reaffirming). The same line of reasoning makes the negative existential statements ("a does not exist") contradictory, for on the Russellian view "a" could not but name an already existing individual the existence of which is then denied by the very statement. The propositional attitude contexts are problematic because one may believe that *a* is F but not that *b* is F, even if *a* is in fact *b*. It is commonly said that the substitutivity of identicals fails in these cases. For instance one may believe that Russell was a philosopher without believing that the junior author of *Principia Mathematica* was a philosopher, even though Russell was the junior author of *Principia Mathematica*. Or, to restrict the examples to proper names only, you may not have a faintest idea of who Saloth Sar was, but know that Pol Pot was a nihilistic dictator; still the two names refer to the same individual. Another problem with the propositional attitude contexts is that one may believe that, say, Santa Claus is fat, but from this belief it does not follow that Santa Claus exists. The rule of the existential generalization fails in these types of cases.

The best straightforward direct-causal attempts at explaining the four puzzles that do not commit some kind of *ad hoc* moves like postulating strict semantic-pragmatic divide are those offered by Michael Devitt. Devitt offers solutions to both types of the existential statements such that make them meaningful, not tautologous and contradictory. According to him the positive existential statements are meaningful because there exists an object on which the causal links underlying the referential expressions are grounded, and the negative existential statement are meaningful because there exists a causal network which is grounded on an "imaginative act" giving rise to a type of referential expression; for instance a story or a hallucination that are, or were, taken to be about a really existing object. I can in fact accept these solutions, provided they can be given a formulation on the informational terms (such will be given in sections 8.1.4 and 8.1.5). However the explanation of the negative existentials is too narrow: the source of the causal chain can be for all intents and purposes a normal act of postulating an individual. This shows, moreover, that the causal links are not the primary factor in these explanations of the existentials even on Devitt's account. Instead it is the kind of information, for the "imaginative" groundings need not differ epistemically at all from the normal groundings of proper names. The consequence is that the names are used *as if* they refer genuinely, i.e. they are used *purportedly* to refer, as are all proper names in the contexts that are about fiction or such topics when that is known to the participants. That is why the negative existential statements are not contradictory: they purport to give us new information about the ontological status of the purported referents of the names in question.

If Devitt's account of the existential statements is provisionally acceptable, the same could not be said about his solution to the identity statements. As pointed out, Devitt recognizes that the direct-causal theory needs some kind of notion of sense as the mode of the presentation of the referent, albeit a non-Fregean one as he emphasizes. His proposal is that that kind of sense consists of the types of designational chains originating from the referent (Devitt 1981: 153, 236), (Devitt 1989: 227-31), (Devitt & Sterelny 1987: 58). The cognitive difference between the identity statements "a is a" and "a is b" is explained by the different designational chains grounded in the same object. In spite of the fact that I said earlier that I categorically refuse to take seriously

any kind of non-cognitive notion of sense let us take a look at Devitt's proposal. His account is in danger of being vacuous. For consider the question how different the causal networks or the designational chains must be for there to be a difference in the sense? In other words, how one singles out the sense-relevant parts of the networks on the purely causal basis? Take for instance the statement "Manoel Francisco dos Santos is Garrincha". This is an informative statement if one did not know the identity before. According to Devitt, now that one knows the identity, the networks underlying "Garrincha" and "Manoel Francisco dos Santos" have "merged" in one's brain that resulting in the informativity of the identity statement. The problem with this solution is that the relevant causal parts of the respective networks *qua* causal could not be individuated to explain this. One needs recourse to the factors that make for the differences in the causal links, and these factors are again informational: as I argued only information keeps the referential links alive so that the names can be used meaningfully to refer. Here I could not help but quote Evans' brusque expression of the problem:

"And it is quite obscure how, if one mental state represents a particular object in virtue of one sort of causal relation to it, and another mental state (of the same subject) represents that object in virtue of another sort of causal relation to it, the sheer differences between the causal relations could generate a difference in *content* between the two mental states, given that it need not in any way impinge on the subject's awareness." (Evans 1982: 83)

In a later paper Devitt admits this point, in fact endorses it (Devitt 1989: 227-8). But now the vacuousness of the solution becomes obvious because Devitt has admitted that the designational chains merge in the brain of the speaker. What precisely happens is that the proximal parts of the chains get registered by being neurally processed in the brain. It is true that the neural processes can be said to be causal but it is essential that they also constitute the information processings that in their turn constitute our cognitive activities, the neurocognitive information we individually possess. Put in modern terms the parts of the causal links *qua* informational in the brain are thus at least very similar to the Fregean senses (*pace* Devitt). So Devitt's solution is not an improvement at all but in effect just the old sense theoretic one.

What about the propositional attitude contexts? Devitt has discussed the matter in earnest (Devitt 1981: ch. 9, 10; Devitt 1989: 229-31). In short he is after a solution to the problem of *exportation*, i.e. under what conditions singular terms appearing in the propositional attitude contexts can be said to refer to a definitely existing individual. When for example the expression "the junior author of *Principia Mathematica*" can be said to refer to an individual (Bertrand Russell) as it is used by a certain speaker when that speaker believes that the junior author of *Principia Mathematica* was a philosopher. (Devitt assumes the descriptions to be singular terms.) Devitt's account says that exportation goes through when the singular term in question is non-empty and designational, i.e. when it has a referent and it is grounded on that referent.

But this is not yet sufficient even by Devitt's own lights – which he does notice. What more is needed is that the speaker "has the referent in mind". As we saw Devitt spells out this notion along the by now familiar lines that there must exist a designational causal chain from the referent to the occasion of the use of the singular term. Moreover

Devitt tries to save Frege's "functionality principle" which he takes to state that the meaning of a statement or a thought depends on the meanings of its components. In regard to the propositional attitudes it follows that the sense of a belief is made up of the "customary senses" of its components. For Devitt these are the causal mechanisms, or parts thereof, by which we refer. If you believe that Cicero was an orator, the senses are the (subsets of the) designational chains underlying the use of "Cicero" and "orator". As Devitt says, "Cicero" and "orator" specify the modes of their reference (Devitt 1989: 238).

The immediate problem with this solution is how the (subsets of the) designational chains are told apart. Obviously not all of the chains grounded in the referent will do, for this does not allow any finer distinctions within them. So we must presuppose that their variance from person to person is what is relevant here (though Devitt's discussion of this seems not quite clear to me (Devitt 1989: 227-8)). But that will not do either because it does not explain the differences in the associated senses of the singular terms. The links in the causal chains are distal, so in the overwhelming majority of cases the speaker has absolutely no clue at all what their courses have been: the major part of the causal network eludes the speaker epistemically. We do have access to the proximal parts and effects of the networks, but our mode of access to them is cognitive and informational (I know that someone is such and such, but I do not know my brain processes, the proximal causal patterns, when I am thinking of that individual). The earlier proximal links do not get into our brains *qua* causal; the only way they could do that is for them to make a cognitive difference for example by carrying a "mark". That is, when they show by themselves that there is a difference in the causal mechanisms that are supposed to constitute the differences in the respective senses. But that could not be done without informational clues about how the links were created and retained, and that requires, to state it once again, information about the referent as well as the courses of the causal links. And that is what the puzzle about the propositional attitude contexts is all about. (As it is also with the identity statements. So it is no wonder why the two puzzles are usually discussed together. And it is no wonder why Devitt's solutions to both of them fail together.)

It also seems that there is a revealing feature (or an error from my point of view) in Devitt's proposed solution to the exportation. He writes that exportation "is in order if and only if 'b' is designational and non-empty" (Devitt 1981: 226). This requires that from i) a person *x* believes that *b* is *F* (where "b" is opaque) and from ii) *b* exists, one can infer that *x* believes *of b*, that it is *F* ("b" is transparent). But this seems to make the whole exportation trivial, for it guarantees that if "b" is a non-empty term (it has a referent, namely *b*) and is designational (is grounded in *b*), exportation goes through *without* the subject needing to have anything to do with this; the subject is not even required to know that *b* exists. Here we can clearly see again the main problem with the direct-causal theory. By claiming that knowledge and beliefs are not necessary (or sufficient) for reference, the way to the solution of the puzzles is simply cut off. Especially the propositional attitude contexts measure the needed amount of information for a reference (exportation) to succeed. If one does not know who the referent of "b" is, one could not export "b". What Devitt says, in effect, is that one does not have to know who or what *b* is for exportation to go through. If this is so, why, then, the problem with the propositional attitudes have become and keeps on being so pressing in the first place? It is a problem and will not go away by trying to answer it with a theory which so cleanses reference that it becomes divorced from what it

actually is: one form of human action, inevitably bringing the epistemic and cognitive paraphernalia with it.

It should also be realized that bringing in the semantic reference instead of the speaker reference, as is the case with Devitt's example, his solution would seem to succeed in most cases. That is because the linguistic, or epistemic, community includes those who know who or what *b* is (the "experts"). But I have already argued that that solution is not wholly cogent. Moreover the propositional attitudes are the attitudes of a particular individual, so the "epistemic state" of the community must be kept out of the considerations.

It should be noted that Devitt takes up some cases that he admits do not yield to his account (Devitt 1981: 271-4). There are three types of them: i) one has an access to a designational chain but we do not think that he has the object in mind nor that exportation is in order; ii) one has no access to a designational chain but seems to have an object in mind and exportation is in order; iii) one has no access and no object in mind, but seems to have a belief about the object. Devitt's discussion of all these cases is quite short, and we do not have to present the particular examples he gives because his answer to all three types of cases is in effect the same. Devitt resorts to intuitions. Although all of the three cases seem to show that his account is in trouble, he says that intuitions about exportation do tend to vary among people, so his theory may emerge unscathed when developed further. Devitt says also that the theory may be able to change the intuitions. Be that as it may the notion of having in mind is the crux of the issue. (And leaving aside the point that if intuitions are allowed to change with one's theory then they can not be used against the sense theory by the direct-causal theorists, especially when the intuitions involved clash with one another, as seems to be the case especially with the propositional attitude contexts.) I think that we can correctly say that the intuitions as to when one can be said to have someone or something in mind are not firm because - once that notion is severed from Devittian causal chains - it is the amount of cognitive information about the referent that determines the reliability of the reference. In general the less information about the referent, or less the central information, the less sure we are about successful reference having been made, hence of having the referent in mind - contrary to appearances. This is not a yes-no problem, but rather exhibits a continuous gradation of cases depending on the amount and quality of the information a subject possesses. Consequently it is no wonder that intuitions fluctuate.

Notes

1. The following quotation, besides that it is quite hilariously ironical in putting Kripke good company with Frege and Russell, is also relevant to the present issue: "Kripke's thesis here is greatly supported, I should think, by the fact that we would not ordinarily even know what was meant by 'fixing the meaning', 'giving a definition', or 'giving a synonym' for a name. It is hard to believe that either Frege or Russell, when they were not being careless, ever actually claimed that descriptions did any of the things." (Linsky 1977: 49) - Searle seems to be after some strong semantic relation between proper names and the descriptions (Searle 1958: 172). If that is so, his cluster theory fails if it is taken also to be a version of the sense theory of reference.

2. This argument occurred to me while I was reading Bar-Elli's paper on Frege (Bar-Elli 1981). That it did is easily understood for in one place Bar-Elli seems to say as much in effect himself: "...but then on Frege's own assumption, that *a* is *b* would itself be known a priori. This, however, is exactly what Frege tries to avoid." (Bar-Elli 1981: 142) I would add that Frege does not try to *avoid* it; rather there is no need to such move on Frege's view.
3. This statement is anachronistic, for the notion of rigid designation is Kripke's own label. Also the opinion that both Frege and Russell held the view "one description per proper name" is empirically false, at least in Frege's case (remember Frege's Aristotle note). Robert May argues (May 2006) that Frege did hold that in any language a name in it has no more than one (complete) sense, i.e. that there is no variability of the senses with respect to a particular name, *prima facie* contrary to the standard interpretation of Frege's Aristotle note. (Thanks to Panu Raatikainen for drawing my attention to this paper.) However May's argument so to say equivocates practically. Here it suffices to state the outline of my would-be detailed argument. The complete sense is a thesis about the ideal scientific language only (as May himself acknowledges), so the vagaries of the natural languages still remain haunting. In other words, as long as we do not possess the ideal language our grasps of the senses of names are bound to remain partial. In a way it is then true that every speaker speaks his own language. Therefore it becomes *very* understandable that the referents and the extensions of the expressions different speakers use (especially those with the same syntactic or orthographic form) have to be the same for communicative purposes. But that suffices for the communicative purposes as well (even if it is not known that the referent is the same). So May is correct but his argument is rather irrelevant – especially because May accepts that a sense may be only partially grasped by different speakers. (To be noted, I endorse May's argument that to the extent we have *conscious* access to the senses – or just to the parts of them (see chapter 6) – they can be characterized only by descriptive means, but from this it does not follow that the senses are wholly descriptive.)
4. So there is also this subclass of the definite descriptions that are rigid, in addition to those considered by Recanati above. It seems that these latter descriptions are rigid because of the nature of their referents, like "the only even prime" shows. This seems to be closely related to the issue of essentialism, unlike the descriptions fixed with the "actuality operator". (Another class of rigid descriptions is made up by the Quinean predicates like "x Socratizes", which express the property of *being Socrates*, thus having Socrates as its sole extension in every possible world. But these kind of rigid descriptions are trivial in relation to the Kripkean argumentative context, as Linsky points out (Linsky 1977: 54).) Note also that actualizing the descriptions and fixing the referent by the descriptions are not equivalent. The description that fixes the referent can be taken as rigid, but any description applied to the referent can be actualized, thus rigidified.
5. In a way Kripke admits this, for he states that in Searle's cluster version "in a rare class of cases, usually initial baptisms" we find all the features (2)-(5) that characterize Searlean version (Kripke 1972: 290). But considering the central role of the baptisms in Kripke's picture, and the allowable range of the namings, is that class so rare at all?
6. It has been countered to me that the referential intuition for the direct-causal view is certainly much more widely held than my contrary intuition. I have two points to make in response. When has the argument from authority/majority become sound? And what has

truth to do with the opinions, or intuitions, held by the majority? I guess that the intuition that the Earth stands still and the Sun revolves around it was as strong and as widely held an intuition as anything is (and still seems to be); but Aristarchus and Copernicus did not accept it, and were correct in not so doing. My answer to the second complaint that rare individual intuitions should not be respected is that it manifests a complete misunderstanding of the role of counterevidence in science (and I assume that in philosophy we ought to respect that, too). The relevant point is that a contrary intuition, if intuitions *are* regarded as evidence, is then an instance of counterevidence, a *datum* that needs to be explained (away). It is not to be dismissed as a rare and maybe isolated instance, but it is a *fact* to be accounted for. My second guess is that many foundational scientific developments would not have seen the light of the day (but the darkness of the night of the authority and tradition) if the contrary facts, seemingly unaccountable phenomena, were just ignored (though many of those *were* ignored for some period).

7. One could also use a shortcut argument and point out that Salmon's notion of reference *simpliciter* is *non sequitur*. For he assumes that there is only one intuitively uniform notion of reference. But there may well be more than one. Indeed, the contrary intuitions indicate that this is so, i.e. that the descriptions associated with a proper name are the theoretical clothing of another notion of reference *simpliciter*. This further indicates that the notions of reference *simpliciter* depend on larger theoretical considerations. In short, in some situations the operative notion of reference could be the Salmonian (causal) one, but in others it is the notion that the sense theory accounts for. In fact my argumentation will have as one of its main results that the situations in which the former notion is "at home" are the demonstrative contexts, but even then only together with the sense theoretic factors; in all other situations the sense theoretic notion of reference is solely operative. (One could try to distinguish the larger considerations as "merely" pragmatic ones. But I think that in that case there seems to be no viable explanatory distinction to be drawn between semantics and pragmatics in relation to the reference research. I praised Recanati's account because it invests quite extensively on the pragmatic factors. I take it that his account can be taken as an argument for the explanatory mesh of semantics and pragmatics.)
8. Michael McKinsey argues that the problem with *causal* intentions to co-refer (intentions not individuated by anything descriptonal), i.e. intentions which are constituted by factors belonging to the causal connections, is that this requires, for the speaker to know that he has that very token intention, that the speaker "must also know something about the chain of events by which this intention was brought about." (McKinsey 1978: 185) The reason is that the knowledge of the very token intention is a necessary feature of the co-referential intention *qua* intentional; the general intention to be co-referring is forceless in relation to the particular (token) intention with a name "NN". But that is generally false: one need not engage in any empirical investigation going beyond his current mental states to know what he intends to. Hence the causal theory's reliance on the intentions fails.
9. So I am in fact denying Devitt's claim that people have false beliefs in many cases about the referents of proper names. But this is surely an empirical issue that could, and should, be decided by research that records people's beliefs about the bearers of the familiar proper names.
10. The relevance of information with respect to the borrowings of reference by the abilities to refer, both by proper names and by the descriptions, is revealed also by what Devitt

says about the *loss* of a particular such ability: "This loss of ability is a failure of memory. For a sense theory, what is forgotten is the required associated descriptions. For the causal theory the inability to produce descriptions usually associated with a name is *evidence* for loss of ability with a name, but does not *constitute* it." (Devitt 1974: 188, n.11) The problem with this view is that the distinction between evidence and constitution is left totally unclear in relation to the role of the descriptions. And what specifically constitutes the loss of an ability to refer in the causal theory? Are there any other relevant memories than the memories about the referent of the name? In other words, if abilities to refer are not constituted by the senses (expressed by the descriptions, and cognitively speaking: beliefs and memories), why even bother to invoke the memories in explaining the loss of particular referential abilities? According to the direct-causal theory the causal links explain having in mind, period; the cognitive stuff is totally irrelevant as evidence.

11. Here I only note the point because it becomes obvious in Devitt's explanation of his example of the Liebknecht case (Devitt 1981: 140-3). As the reader recalls ignorant Joe catches the name "Liebknecht" upon the conversation by his friends. But being ignorant he spreads a referential tangle as to whether it is Wilhelm or Karl Liebknecht to whom his uses of "Liebknecht" refer, for there seems to be no fact of the matter to tell them apart (by the causal account). Devitt invokes Field's notion of partial reference: "Liebknecht" refers partially to Wilhelm and partially to Karl. But that explanation does not differ from what the sense theory of reference delivers, for there are no individuating descriptorial means on Joe's disposal by the help of which it could be found out whether Joe refers to Wilhelm or to Karl. The causal account here rides a piggyback on the sense account. This point generalizes to other cases where the routes of the causal links are not known and information related to the bearers of the names are of no help either.
12. In this I find myself in company with Salmon who states that equating Fregean senses with "...the linguistic network or a chain that secures the referent is...on the order of category mistake." (Salmon 1986: 70-1)
13. Another use of proper names is to "verb" them, like in "I'll Sarah Young you, you #□%&!?", as one once teased me. I take it that the relevant elements of the meaning of "SY" on which the tease was based are obvious (if one knows about SY, that is).
14. Though there are also strong contrary intuitions to the causal ones: we do seem to talk about future individuals with descriptorial specifications ("the first born baby in Finland in the year 2010"). We can also talk about individual black holes in the centers of particular galaxies, about the heaviest whale in the oceans, and so on.
15. NN = *notetur nomen* or *nomen nescio*; meaning "note the name" or "name not known". There are well known arguments against this view by Kripke (Kripke 1972: 283-4) but I concur with those who favour the explication and take Kripke's arguments to be weak.
16. I mention Piaget here to anchor the background of my proposal (and for historical accuracy). That does not mean that I accept his overall theory; at least I prefer to stay uncommitted with respect to his strong antinativism.
17. It seems that the basic schema helps to explain also Donnellan's notion of referential use of descriptions, at least in those cases where the intended referent is in sight (so that it can be pointed to, if needed). In these situations the basic schema takes over the referential

role; the specific information the description conveys cannot, so to say, resist in these essentially perceptual situations (even if it is correct).

18. The basic schema may also help to explain what Evans calls "generality constraint". The generality constraint assumes that thinking involves conceptual abilities (in contrast to mere manipulation of distinct symbols) and that thoughts are essentially structured (Evans 1982: 101-5). The substance of the generality constraint is that if a cognitive agent is able to think that *a* is F, he is also able to think that *a* is G, *a* is H and so on. But he is also able to think that *b* is F, *c* is F and so on. In other words there are no isolated conceptual skills with respect to the objects of thoughts and their properties in the mental life of any cognitive agent. In this way the generality constraint partly characterizes what is to be a cognitive agent. Now the formal presentation (*a* is F) displays it admirably how the generality constraint works at the fundamental level of enabling cognitive agents to think about anything. So I propose that the basic schema of reference explains this ubiquitous feature of the cognitive agents being able to think subjects being F, G and so on, that is involved in any thought which is about someone or something. - Evans denies that the generality constraint applies to the information processing in the brain (in contrast to *person's* thinking), i.e. he thinks that the brain's processing modules are distinct (like the modules processing the speed of sound and the speed of moving objects) (Evans 1982: 104, n.22). But as an argument it fails. For all we currently know there could be such information processing systems in the brain that execute just the general ability the generality constraint focuses on. The basic schema of reference could indeed be such a general system.
19. It seems that the idealizing tendency, which goes with the tendency to take the causal connections to be without any real problems, is clear in Devitt's presentation of his theory (Devitt 1981). Although he reminds us on a few occasions that there are difficulties with the details of his theory, and discusses them also, the tone throughout his book is that the Fregean and the description theoretic factors are referentially almost epiphenomenal; only the causal connections, the d-chains, count in explaining the abilities to refer, the reference borrowings, non-Fregean senses of proper names, the different conceptual roles of them, the justifications of exporting singular terms in the propositional attitude contexts, and so on. In reading the book one rather soon gets the impression that the causal links work magic and one is left wondering how anyone ever thought that the Fregean sense and the description theories could provide viable explanations of reference of proper names at all. But Devitt's theory is the most thoroughly developed causal account of the reference of proper names, so the impression immediately suggests a general *tollens* argument: the direct-causal theory of reference just will not do.
20. And not so tacitly sometimes. For instance David Kaplan, one of the proponents of the direct-causal theory, asks in one occasion that, if pointing can be a form of description, why could not description be a form of *demonstration* (Kaplan 1990: 24). (Look also at what Kaplan says about the relationship of the demonstratives to proper names in sections 6 and to the language learning in section 10 at the same paper.)
21. Since I first read about the Twin Earth scenario and the related arguments against the Fregean theory, and for externalism and wide content/meaning, it has always seemed obvious to me that they could not be correct because they ignore what is known and not known in the scenarios by the subjects. It is, at last, reassuring that there is at least one other who seems to think precisely that way too: Frank Jackson points out that the neo-

Russellians (though he does not name them) are in troubles because they must help themselves to the distinction between information conveyed and knowingly conveyed. But then "...we have the problem of giving an account of what is knowingly conveyed." (Jackson 2004: 262). In other words, and generalizing, the neo-Russellians and direct referentialists, and those who use the Twin Earth scenarios against the Fregean views, do – and have to – presuppose an epistemic element to the peril of their core claim (that the object of thought with cognitively differing expressions both referring to it is the same in every case and at all times, i.e. the referent itself).

22. To pre-empt the issue of begging semantic and pragmatic questions before the section 4.4 where they could arise anew I want to bring up three subjects here. 1) Salmon's own information-semantic framework begs several questions, for example he takes information to be very strong metaphysical notion (it is eternal and same for every thinker and speaker with respect to every distinct declarative sentence). But the most serious failure of his framework is internal. In his eagerness to prevent any Fregean factors from creeping in he postulates *guises* under which objects are thought. But that is clearly equivalent to the postulation of the Fregean senses for all explanatory purposes. (Graeme Forbes has observed that also (Forbes 1989: 135, n. 19).) But what is more his formal framework, with its "contours", "schedules", "characters" and such factors contains also *intensions* which give the extensions in any possible world (Salmon 1986: 32-4). How is Salmon to avoid that those intensions contain the Fregean factors that determine the referents? It strongly seems that Salmon's neo-Russellian semantic framework works only because it has the central elements of the Fregean theory tacitly built into it. (This charge becomes even more pressing when it is noted (see note 12) that Salmon rejects any causal links between objects and expressions as having anything to do with information (Salmon 1986: 70-1).) – 2) Salmon also charges the Fregeans of begging the question in that they do not differentiate between semantically encoded information and pragmatically "imparted" information by a sentence. But Salmon is led in effect to a stipulative mode; he is forced to keep the semantic information so narrow that it only contains the referents including properties and relations (in general, the information values) of the expressions involved (this is what his neo-Russellian "doubly modified naive theory" of information content claims, in fact). The stipulativeness in Salmon's approach becomes even more evident when he says, in countering the suggestion that his *guises* are in fact Fregean notions, that whatever the *guises* are they are "entirely separable from the semantic nature of the relevant sentence..." (Salmon 1986: 120). Salmon differentiates between semantics proper and pragmatics; the latter deals with the way "a thinker is familiar with the thought content [a singular proposition]..." (Salmon 1986: 174, n. 2). This maneuver, when it is coupled to the drawbacks of that theory (for example that the *guises* are operatively Fregean), justifies the application of *modus tollens*. But in a way I am quite willing to let Salmon to have his austere conception of semantics, for I suspect that ultimately it will coil into itself without any explanatory power with respect to the referential *communication*: the singular propositions are rather uninteresting because they can be had only after our epistemic work is done, i.e. when we eventually become to know the ontological inventory of our world - until then all the claims to the individuals being the constituents of the singular propositions are hostages to the temporary states of knowledge. -3) One particular result of this that I favour is to break the distinction between the encoded and imparted information. Whether information is semantic or (merely) pragmatic does not count. What counts is what explains what, and the modified naive theory loses in that game, with the *guises* and all. I think that this morale generalizes: any attempts to draw a wedge between semantics proper and pragmatics are

bound to be stipulatory and without explanatory force when the subjects are reference and thinking as forms of human action. (As already mentioned an enlightening antidote against the overly "semanticizing" attitude is (Recanati 1993).) By pragmatics I here understand factors that have to do with the speaker's and the hearer's contexts, their perceptual surroundings as well as their beliefs (this conforms closely to Recanati's view of pragmatics). In contrast "wide pragmatics" may be barren with respect to questions in semantics and "narrow pragmatics". The "wide pragmatics" deals with Gricean maxims and such factors. (Therefore it could not be held against me that with my postulation of the metaprinciple in section 4.3.2 - which says that it is tacitly acknowledged by all competent speakers that if there are in fact two distinct individuals for the same proper name, that could make differences in the referential practices and in the senses, *but* that that is irrelevant as long as there is no indication that such is the case in fact, i.e. that they are two different referents –Am I here ignoring my own view of the non-existence of any strict explanatory boundaries between semantics and pragmatics? I am not because the metaprinciple belongs to wide pragmatics.) -3) There is the puzzling account when Salmon states that according to the Fregean view, when someone believes that there are two distinct individuals bearing the (syntactically) same name when in fact there is only one self-same individual, the name must be as univocal as it is to those who know that there is only one individual concerned. But why are we supposed to be forced to adopt *our* epistemic perspective as exclusively applicable to the subject? We are trying to characterize the innocent speaker's perspective on things, and we can do that by keeping his senses separate (i.e. we acknowledge that he thinks there are two distinct individuals). We do *not* "use the name to refer to *our* sense" only (Salmon 1986: 171). Of course we must use our senses when talking about the speaker's senses, but we can characterize them being separate from each other and as giving at least a sufficient approximation to the speaker's senses. If we could not, ascribing specific propositional attitudes to anyone who has not our epistemic perspective on things would be futile, but it is not. Salmon has made a singular mistake here, not the Fregeans. (It seems that he is here the victim of the urge to adopt the epistemic god's eye point of view of which I warn about because it is bound to lead to distorting results in semantics.)

23. Maybe a more realistic example could be given with names "Gaurisanker" and "Qomolangma" which, I believe, both refer to Mount Everest.
24. Russell held also this view, albeit – of course – expressed in terms of descriptions: "The description required to express the thought will vary for different people, or for the same person at different times. The only thing constant (so long as the name is rightly used) is the object to which the name applies." (Russell 1912: 29-30) Scott Soames uses the notion of information asserted by a sentence and argues that no substantial descriptive proposition (in contrast to singular proposition) is expressed by any such assertive utterance because different speakers know different things about the referent of the proper name used in the sentence (Soames 2002: 64-5). But as I already argued Soames' fallacy lies in the fact that the sense theorist (even if it is assumed that the senses could be expressed descriptively without a remainder) is not committed to such common-information-to-all-speakers demand: remember that it is part of the sense theory that speakers have different information, i.e. different senses, about the referents of proper names. I would like to ask bluntly *how* could two people *ever* believe *exactly* the same thing? Our experiences and information we individually possess, hence the senses, differ from one another. So it is extremely improbable that there is any intersubjectively same thing believed by any two thinkers (and by any single thinker throughout his whole

cognitive life). Sometimes it is strongly maintained that if the contents of particular beliefs were not the same, we could not comprehend one another. (I have even met explicit exclamations in the literature that it would be *absurd* if that were not so - but I will not disclose any names here.) Be that as it may, for now, I can not help believing that those who impose this *a priori* demand from the sameness of the content of the objects of the propositional attitudes, in the manner just alluded, are simply confusing the minimal (Fregean) requirement of the sameness of the referent, or extension, with the sameness of the whole content of a particular propositional attitude, i.e. the total sameness of the senses related to a proper name. (When one meets *a priori* demands, generally one should be ready to fire the charge of question begging. Here it is that the proponents of the sameness of the content beg the question against the similarity of the content of the objects of the propositional attitudes.)

25. It should be pointed out that transworld identity of individuals is a heavy metaphysical issue. For instance one could argue that since Kripke *stipulates* possible worlds, via the actual individuals, he simply begs the question against any different view. A particular different view is the Lewisian possibilist approach that postulates that there exist (many) other possible worlds that are just like our world in their basic physical properties (but that there are also many others which are not, i.e. in them even different laws of nature reign). This approach comes with the counterpart metaphysics of individuals. According to that metaphysics no inhabitant of any world is strictly identical with any actual individual, or with any individual in any other possible world; they are at most counterparts of one another (one counterpart per world, though this may not be necessary) depending on the (different) similarity considerations of their properties. However the issue is not so clear cut. For the Kripkeans may argue that it is precisely the similarity relations that beg the question against their view. Another way to put this is to note that the Lewisian metaphysics does not countenance rigid designators at all (because there are no strict transworld identities). In this work I have been accepting for the sake of the argument the notion of rigidity, but as comes to the metaphysics of individuals as such, distinct from the issues in the philosophy of language and semantics, I side with Lewis. In particular I cannot help but wonder why there are, in Lewis' words, "incredulous stares" the view has been facing. These stares should be dimming. My reason is that the possibilist view is currently a scientifically justifiable option. That is because the multiversum hypothesis is gaining ground among the cosmologists and astrophysicists. That our universum, our possible world, is only one among many others (most likely different from our world in different degrees with respect to the basic physical properties and the laws of nature) seems to be the best explanation of, for example, the fact that in our world life as we know it has been able to evolve (although the margin of error for that happening is very thin indeed in terms of the values of many basic physical constants and parameters). In other words the issue is also as much an empirical one as it is metaphysical. – I would like to add that when the counterpart theory is wedded to the Lewisian possibilism, the categorical denial that any two individuals are strictly identical across any two possible worlds might not be true. For if our universum is totally deterministic, then another possible world that has the very same origin (with the same basic physical laws of nature and the same distribution of matter/energy) evolves as an exact copy of our world. Thus in it every individual is a totally identical counterpart to a corresponding individual in our universum, and the counterpart mapping is strictly 1-1, i.e. the worlds are perfectly identical. Consequently it is not wholly true to state that the counterpart theory eschews rigid designators, for with respect to the possible world just envisaged all the names of the

actual individuals are rigid; hence at least the letter of rigid designation is respected in the cosmologically supported Lewisian counterpart theory.

26. With Salmon we have a case of theoretical blindsight on this issue. For he says on another related occasion that "Strictly speaking, each of the argument applies to a proper name or indexical singular term *as it is in a possible context*. This is important...also for proper names such as 'Salmon', which may be the name for two or more individuals at once." (Salmon 1982: 24, n.26) So he sees the need for the actual identification of the referent of a proper name, i.e. the problem of the *disambiguation of proper names*. Salmon's lapse is even more curious because the contextual relativity he mentions is just what his caveat from the relativization of a use of a sentence states. In this connection I should mention also what Devitt and Sterelny say: "We do not overlook the importance the linguistic and non-linguistic *context* of an utterance but see all aspects of it, aside from the relevant parts of the causal theory, as having only evidential significance. The context guides the audience in removing the ambiguity; it supplies evidence of what the speaker has in mind and hence evidence of the semantic reality, but it is not that reality." (Devitt & Sterelny 1987: 59) But the very distinction, in relation to the individuation of the referent, between reference proper and evidence for a specific reference, is problematical. Once the causal theory is shown to be explanatorily vicarious with respect to the Fregean theory what is left are the "evidential" factors, and these are informational and so, to emphasize it again, the sense theoretic factors.
27. Soames discusses disambiguous proper names (Soames 2002: 96-103). He answers the problem of the individuation of the referent, or the determination of the semantic content of the ambiguous proper name in a context, by arguing that the sufficient condition for that is that a particular name used has the same referent in all contexts of its use (Soames 2002: 101). This is spelled out, in part, by the requirement that the speaker has acquired a "referential intention" that determines a certain individual as the referent of the name (for example by picking up the name from others and intending to use it to refer to whom they use it to refer). But this method of individuation, i.e. disambiguation, fails as I already argued: the intentions are not sufficiently powerful to determine the referent without related individuating information. And Soames disregards the associated information *in that role*, but his argument fails also there, for it assumes that the information must be *uniformly* shared by the speakers; the sense theory does not require that. Soames' take on the issue here manifests the status of the current common opinion, the received view: it has become uncritically accepted as involving only but the barest of outlines (like the intentions to co-refer allegedly being sufficient for referent retention). Moreover there seems to be a plain counterexample to the soundness of the intention strategy. For what are we to say about the possible case when the causal contacts of the names the current speakers use are to two (or even more) distinct individuals with the same name, but we do not know that and the causal links have become mixed? It could even happen that this is a case with deliberate confusing: the two characters – identical twins, perhaps – have just been amusing themselves in appearing to others in different occasions. What singular proposition is expressed in that case? Is it one that involves both characters? But that seems *not* to conform to the normal practice of thought expression, and the expression of the semantic content of the propositional attitudes in general. We would be taking ourselves to be expressing unproblematic singular propositions (NN is so-and-so), but it seems that the thoughts expressed would involve both NN-characters. Or if not, then what would they express?

28. Metaphysically-minded direct-causal theorists could still argue that the individuation of the referent is ultimately not to be decided on such rather slender empirical grounds as the causal connections, or the Fregean senses, but on much firmer grounds. This leaves them two options (as far as I can see): 1) *haecceitas*, or primitive thisnesses, or 2) essentialism with respect to origin. The first view is argued for by Robert Adams (though not in relation to the issue of reference) (Adams 1979), and the second view is Kripke's (Kripke 1972). I will not argue against these views here in any detail. I really cannot make metaphysical sense of (primitive) thisnesses so I will leave Adams' account aside. Kripke's version of essentialism with respect to individuals I reject because the origin of a biological entity, its "parent" DNAs, could have been different without that affecting differentially in any way his, hers or its later development so that the result would be the same individual as actually. The reason is that there are degeneracies of amino acids, i.e. the relationships between some nucleotides and amino acids are many-one. Likewise the relationships between some amino acids and the resulting proteins are such functionally. Some proteins are still able to function even if some parts of them are lacking or weakened (as in mutations). So the question is what is the relevant range of variability and differences in DNA to preserve the identity of a biological individual? If it is stipulated that only when the DNA is exactly the same is it metaphysically legitimate to talk about the essence of origin as the essence of the individuals, the Kripkean view may still not stand a closer scrutiny. For what are we to say when DNA-identical individuals become so different that all other individuating criteria than DNA give the verdict that they are different individuals? It seems to me that in that case we should at least reconsider the soundness of the stipulation. (By the way should not exact DNA-identical twins be the very same individual on that account, only numerically distinct?)
29. Kripke says about the evidential role of intuitions: "Of course, some philosophers think that something's having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don't know in a way what more conclusive evidence one can have about anything, ultimately speaking." (Kripke 1972: 265-6) Even if one ignores the rhetoric impact this presents a very tall order - so tall in fact that I am stupefied. There is no attempt to explicate intuitions; do intuitions serve as evidence for any issue whatsoever?; why are intuitions the ultimate evidential base in philosophical matters?; what about conflicts of intuitions, especially equally "strong"?; or is the number of those holding the same intuitions (and, perhaps, with equal strength) to be counted only as evidential, the minority intuitions not? In short, this is not critical philosophy but a sample of not a very reflective philosophical methodology. I would say that the more a discipline resorts to intuitions, the more immature it is if it has scientific aspirations. (As I already pointed out, if uniform intuitions were to be respected as evidence, and respected the more so because they were held by the majority of people, Aristarchus and Copernicus would not have proposed that Earth revolves around the Sun; but it does, so even very firm perceptually-based intuitions as evidence can be discarded. And mind you, the firmness in the direct-causal theory seems quite shallow in comparison.) When this skeptical view is coupled to some other recent writings - (Rosenberg 1994), (Hintikka 1999), (Cappelen & Winblad 1999) - that also cast a serious doubt on the cogency of the evidential status of intuitions. The case against them deserves serious consideration. And I could not resist quoting Hintikka here, for his rather harsh verdict on the unargued status of any justification for the use of intuitions is to the point: "The blind faith [of appealing to intuitions] is below the intellectual dignity of philosophers for whom unexamined intuitions should not be worth intuiting." (Hintikka

1999: 130) The only naturalistically respectable attempt at justifying the use of intuitions I know of is Devitt's paper (Devitt 1994). But he seems to give intuitions a role only in the starting stages of semantic research, i.e. in the preliminary carving and characterization of the domain. Clearly, then, the results of such applications of intuitions could be discarded at the later stages. For a scrutinized account of the role and the nature of intuitions in science, with obvious implications to philosophical methodology, see (Fishbein 1987).

30. It should be noted that Almog's view is for unification of referential semantics and cognitive semantics, but only terminologically. He wants to achieve this unholy alliance by subordinating cognitive to referential: "It is *because* our cognition is object-bound and content-free that our semantics is, in turn, purely referential." (Almog 2005: 531) This is but verbal unification for the "content-free" means anti-Fregean. But, as I argue, *because* reference depends on cognitive, on contentful thoughts, and *because* the anti-Fregean arguments fail, no unification could be achieved – only subordination the other way around.

5. ACTUALIZED DESCRIPTIONS

Background

My focus has been on the Fregean theory of sense and reference, although I have pointed out in the appropriate places when my arguments work also in favour of the description theory. But maybe I have been more assuming than explicit in that what is said in defence of the Fregean theory, and against the direct-causal theory, is to be readily understood to apply to the description theory. To this class belong 1) the misconception that the description theory of reference is committed to the stringent semantic equivalences like synonymies and/or analyticities between proper names and the descriptions related to them; 2) the arguments from ignorance and error in that i) the Feynman and Gödel examples are basically beside the point because they invoke descriptions that are not sufficiently individuating to begin with, and ii) they use deference as the mechanism of reference in the way that the description theorists are not forbidden to use also; 3) the arguments for the epistemic symmetry principle and against the relevance of the Twin Earth arguments work for the description theory as well. However, and as Panu Raatikainen has more than once emphasized to me, even if the Fregean theory emerges unscathed from the direct-causal attack, some of the arguments given by the latter theorists may still be effective against the description theory. It is this dialectical situation that forms the background of the argumentation in this chapter. Still, I want to point out that I am not a proponent of the description theory; I endorse the Fregean view modernized. The main reason for this is that the senses are not totally articulatable by any linguistic means. (I say more about this in the following three chapters.) If, as it seems, the description theory gets also defended along with my arguments for the Fregean view (and against the direct-causal view), that is fine by me. The following argumentation is motivated by the inherent interest I find in it and because it seems to me that the points I will be making are new to some extent. And obviously, what suspicions there might be left about the Fregean theory because of the modal argument, the argument in this chapter should put them to rest because actualized descriptions express “actualized senses” also.

The actualized descriptions view

The original version of the modal argument is based on the presupposition of synonymies between proper names and the definite descriptions related to them (as their meanings). We saw that this presupposition has no justification with respect to the Fregean, and neither the description, theory. But the argument can easily be stripped off of that presupposition. Indeed it seems to me that there is an ambiguity with respect to the argument: it has also been advocated without any (at least explicit) recognition of the synonymy demands. I propose to accept that line. It seems to me that the modal argument can be maintained also with respect to the semantic relations of the definite descriptions and proper names outside any stringent theory of semantic equivalences

Let us begin with the statements

- 1) Aristotle might not have been Aristotle
- 2) The most famous pupil of Plato might not have been the most famous pupil of Plato.¹

The first is true, and necessarily so according to the rigid designator account: Aristotle was Aristotle, the self-identical man, so he could not have been anyone else. Apart from some scruples from the metaphysics of essentialism, we may accept this account. The statement 2) is ambiguous – or it has two readings, the so-called narrow and wide scope readings. The narrow scope reading focuses on the content of the *description* as singling out the referent. Therefore the most famous pupil of Plato could not but be the most famous pupil of Plato, whoever that individual is in different possible worlds. That is why the statement 2) is contradictory in the narrow reading. But the wide scope reading focuses on the *individual* and takes the content of the description to have a secondary role after it has picked out the individual. Therefore the most famous pupil of Plato, the very actual Aristotle, might not have been the most famous pupil of Plato; he might have gone to *pankraton* instead.

Kripke used this kind of statement pairs in arguing that proper names and the definite descriptions differ in their respective modal profiles (Kripke 1972). Aristotle could not but have been Aristotle, but the most famous pupil of Plato could have been something else than the most famous pupil of Plato. Consequently, and generalizing, proper names differ from the definite descriptions with respect to their modal profiles. Therefore proper names are not synonymous to the definite descriptions: the descriptions do not give the *meanings* of the names - for if they did, no such differences could arise. But we agreed to drop the synonymy demand. Still it is correct observation that Kripke focused only on the predicative, or narrow scope reading of the definite descriptions. But the predicative use may or may not be the more frequent one in the everyday practice of referential language use. We often say things like “NN did so-and-so” and “NN is such-and-such”, but we also use the descriptions only and say things like “the so-and-so ...”. The former statements, the referential reading, brings the description to the place of the subject in the sentence (or makes it the “logical subject”) just like in the statement 2) above. But Kripke has taken the referential reading and interpreted it to be the predicative reading, i.e. the narrow reading instead of the wide reading, the one that amounts to the actualized description view. This is especially clear when presenting Russell’s theory Kripke plainly ignores the real referentially relevant impact of the context and treats the description related to the proper name via its predicative content only: he takes the correct reading of “Aristotle was fond of dogs” to be the predicative reading of “Exactly one person was last among the great philosophers of antiquity, and any such person was fond of dogs” (Kripke 1980: 7). This begs the issue by ignoring the possibility that Russell’s theory could be made to cohere, or even is already coherent with the actualized descriptions view.² And “in a context” what presumably would exclude other individuative factors – amounting to Kripke’s notion of “conventional use” of the proper name in question – such that it is made amply clear that *only* the actual referent of the description is intended? That is, these other factors need not be given *descriptively* at that context. In fact I would suggest that they could be of whatever kind as long as the actual referent gets singled out. (And they are, as we saw in section 4.3.5.) But anyway because there is this alternative reading Kripke’s argument could not be accepted as it stands. Or at least that is the claim of the actualized descriptions view. According to it we can take the definite description to refer (or to denote; I do not think that the particular term for being about is relevant here) to the actual individual the description describes and then keep that individual as the referent in all contexts, the modal ones included. The upshot of this maneuver is that the description has become rigidified and so to possess

as strong a referential force as the proper name it is related to. This actualized descriptions view has been proposed by Plantinga (1978) and defended by Yu (1980) and Stanley (1997), to mention a few. (See also (Davies & Humberstone (1980: 29, n. 42) who state that Kripke's argument will not succeed against the actualized descriptions view.)^{3,4}

As we saw in section 2.6 Searle's theory could have been more precisely formulated. However he states correctly that we do not *define* proper names. I take it that this statement would preclude such arguments against his view that depend on the demands of semantic equivalences like synonymies and/or analyticities between proper names and the descriptions *if* there were not another place where Searle is explicit that analyticities are involved (see the quotation in section 4.1). So it seems that we could not have a sufficient defence of Searle's view such that could rescue it from the modal argument. The case against him is more conclusive because he also states that if not a single one of the descriptions commonly attributed to a referent, say "Aristotle", were true of the purported bearer of that name, then Aristotle would not have existed (or did not exist). Now Kripke's modal argument – when read even as independent on the demand for strong semantic equivalences – shows that Searle's view is not correct. For it is meaningful to say, i.e. it is possible, that Aristotle existed even if we know nothing about him.⁵ This is also the crux of the "argument from intuition" that we saw above is the last straw in the semantical argument. But these two arguments should not be equated even here. The reason is that the modal argument does not say, or presuppose, anything about the causal links, whereas the semantical argument ultimately invokes them. The modal argument focuses on the modal profiles operative in the way we use proper names versus the definite descriptions. As Salmon explains it the semantical argument contains a modal element in that we are to envision a situation in which the use of a proper name derives from some other individual than the one to whom the description applies (Salmon 1982). In other words in this case the referent is not what it actually is; but in the modal argument the referent is the actual referent of the proper name in question. Therefore only the modal argument is relevant here; only the modal argument purports to support the thesis that proper names are rigid designators.

This issue has sometimes been approached by invoking referential intuitions: it has been taken as understood that it is intuitively obvious that proper names differ from the definite descriptions in just the way Kripke's modal argument shows. But I have to deny the correctness of this approach – for two related reasons, an intuitive and a factual. *My* intuitions support the actualized reading of the definite descriptions: the general schema "it might have been that the F is not the F" has never had an odd ring to it for me. Of course I have known that if the definite description "the F" is taken narrowly, or predicatively as restricted to its descriptive content, the statement is semantically incoherent. But that reading has always seemed to me to be beside the point. I think that this clearly indicates that I have been reading "the F" as an actualized description in relation to its referential force. Assuming that I am not an atypical speaker (in this respect at least) I surmise that this is the case with many, or most, other speakers, so that the actualized reading is the natural one. (If one has schooled oneself with the Kripkean modal argument, his intuitions might have changed so that he only, or first, reads the statement predicatively.) And this is the factual reason: we *do* use the definite descriptions in many cases as referring to the actual individuals they describe individuately. This is more important still when the parity is observed: Kripke has always emphasized that the focus should be on the

ordinary ways we use proper names and the descriptions. Now that ordinary way is, I claim, the actualized reading of the definite descriptions, not the narrow-predicative reading that Kripke's modal argument exclusively relies on. Therefore the modal argument in fact does not have the ground it has been assumed to possess by the direct-causal theorists (and the anti-Fregeans and the anti-descriptivists in general). In other words the modal argument is *non sequitur* for it ignores the wide-referential reading, i.e. the actualized descriptions view.

But my intuitions and factual proposals related to them aside, are there firm *evidence* for the actualized descriptions view as being the primary mode in the ordinary uses of the definite descriptions? Indeed there is, and quite amply in fact. So let me give just few cases which I hope show the primacy of the actualized reading over the predicative reading. When one takes a look at any quiz where the riddle is about an individual the questions help to single out, the role of the definite descriptions as being about actual individuals could not be more obvious. My recent encounter was "Who author?", and the 5 points clue was "Her maiden name was Miller. She wrote in English and used also a pseudonym "Mary Westmacott".⁶ As to non-quiz examples I once heard it asked "did he become the most famous pupil of Plato at those times, or only much later like in the Middle Ages?" during a discussion of the very issues we are focusing on here. No one, obviously, took it that "the most famous pupil of Plato" referred to any other individual than *the* Aristotle we all know (by descriptively expressed information). Sometimes we also use the definite descriptions "out of the blue" in that the proper name of the individual has not surfaced explicitly in the context, like when someone says "the president of France seems to suffer from Aspergerian inferiority feeling", and the audience is currently likely to know that the referent is Nicolas Sarkozy. I do not think that any more examples are needed, for I am sure that if one begins to recall similar happenings the actualized force behind the descriptions becomes evident; *ponderanda sunt testimonia, non numeranda*. (It should be noticed that to try to counter the argument by stating that "of course we use the definite descriptions also referentially the way the above examples manifest, but that is almost trivial, hence carries no argumentative weight" fails. The reason is that this "triviality" just is my point, and it strengthens my case even further: because of it the focus of the modal argument on the predicative reading, and use of the definite descriptions, is even more misdirected.)

So we see that the actualized descriptions view remains unscathed from the modal argument because the referential force of the actualized definite descriptions in actual use is the same as that of proper names. However the referential uses of the definite descriptions, with the tacit actualization-rigidification of them through that use, puts the *original* version of the actualized descriptions view in a sharper focus. It has always seemed to me to be quite an artificial view for it has been taken to involve the explicit fixing of the "actuality operator" to the descriptions. It is thought that the adequate way to understand the actualized descriptions is to present them by statements schemas like "the *actual* such-and-such is so-and-so". But the evidence concerning the normal referential use of the definite descriptions we reviewed above shows that the "actualization operator" is *implicit* in most cases of the use of those descriptions.⁷ That is, it is tacitly understood that the purported referent of the definite description used *is* the actual individual; its actuality needs not be distinctly emphasized in any way. Consequently I propose that we should get rid of the artificiality of the original actualized descriptions view by taking the "actuality

operator” to be implicit in the use of the descriptions. In other words, the actualized descriptions view is not just a syntactico-semantic theory designed to rescue the description theory from the modal argument. We should instead see it as what it is: an expression of the central feature of the normal referential use of the definite descriptions. The original actualized descriptions view is then also seen to have been on the right track but for a wrong reason. It was based on the shared belief with the direct-causal theorists that the only proper way to understand the definite descriptions is the predicative reading. (Apart from the Donnellanian cases, which are mostly demonstrative (see section 8.2).) Because of that the explicit imposition of the “actuality operator” is bound to be an obvious move. But that move is the result of an inadequate understanding of the normal referential use of the definite descriptions. The description theorists have so to say succumbed to Kripke’s maneuver.⁸

When it is recalled what I argued in section 4.5, anyone who accepts that proper names are rigid designators should not recoil from accepting the referentially used actualized definite descriptions as rigid also. The rigidifying of proper names involves the same tacit application of the “actuality operator” (or perhaps: index) as do the referential actualized uses of the definite descriptions. If one mistakenly stares only at the explicit descriptive content of the descriptions one is bound to fail to observe this rather obvious fact. It seems to me that the deeper explanation of this mistake is that the functional differences between proper names and the definite descriptions have not been made too clear. A proper name refers to an object – when used to refer, i.e. not used predicatively – and names it. But spelling out what object the name specifically refers to requires descriptions (expressive of the senses) because of the lurking problem with the disambiguation (relativization of use). In *that* function the descriptions function predicatively. But when a definite description is used to refer to that object, the object it also describes, it assumes the same referential force as the proper name it is associated with. If there are several other descriptions associated with the proper name they function predicatively, i.e. they support the referentially used “primary description”, just as they support the proper name in whole when only the latter is used. (The mechanism of this will be further elaborated in the next chapter.)

Soames’ argumentation

Scott Soames has recently argued that the actualized descriptions view could not be sustained (Soames 2002: 39-50).⁹ His argument invokes the interplay of modal considerations with propositional attitudes. He amasses six arguments to show this. I will deal with all of them in turn. The first argument is that the rigidified-actualized descriptions, unlike proper names, do not refer to the individuals they specify in the possible worlds (counterfactual situations) where those individuals do not exist. However Soames himself counters this argument by admitting that it depends on the assumption that the rigidified descriptions make existence claims (Soames 2002: 326, n. 33), but that need not be assumed. The second argument is the same as the epistemological argument. Soames formulates it thus: the proposition expressed by ***if n exists, then n is D*** is not knowable *a priori*, but ***if D exists, then D is D*** is so knowable (Soames 2002: 41). Soames claims that this difference holds also when the descriptions are actualized-rigidified. But we need not dwell on this argument any more because it was countered in section 4.1 by showing that neither the Fregean theory nor the description theory is committed to, and need not be committed to, *a priori* epistemological claims about the senses and the descriptions. Moreover

Soames' formulation seems to flout the distinction between the speaker and the semantic reference (or comprehension). If a speaker relates to *n* the description *D* it is *a priori* knowable (indeed known by her) that if *n exists*, then *n is D*. Someone else may relate some other description to *n* so it is not known *a priori* by him that *n is D*. But if the semantic reference is assumed, and in the idealized form that the descriptions express all that is known, say, about Aristotle, then the statement that "**Aristotle is D**" ("**D**" including all the descriptions) is known *a priori*. But the semantic reference perspective need not be assumed in the first place.

Soames' third argument invokes the semantical argument. But that argument was also shown to fail on all of its crucial aspects in chapter 4, so no more about it here. Let us then turn to the fourth argument which invokes Donnellan's observation that proper names as compared with the rigidified definite descriptions are exportable from the scope of the propositional attitude verbs without problems – or at least without such problems that the descriptions face (Soames 2002: 42-3). (Soames mentions here (Donnellan 1979: 54-5).) But this presupposes, as Soames notices, that knowledge that *n is F* is always *de re* of the referent of "*n*", but that is not always so with the descriptions (*D is F*). Consequently synonymy between proper names and the descriptions fails.¹⁰ In addition to the by now out-moded assumption that the definite descriptions should be synonymous with proper names one can ask why is knowledge that *n is F* always supposed to be *de re*? On the face of it there is no compelling reason for that assumption. For example I do not know *de re* to what "Butcho" refers in the sentence "Butcho is fat". (I assume that for *de re* contact, hence knowledge, it does not suffice that there is *some* causal relation from the referent to me. That would make *de re* demand here all too easy to fulfill – even if there were not the insurmountable problems with it that there are, as we saw in chapter 4. I will also argue in section 8.1.6 that *de re* is *not* always tied to proper names only.) And I do not know that because I do not know who is Butcho (or if he does exist); I only overheard this utterance. Moreover from the way Soames himself explicates *de re* beliefs about the actual world – "...that is, believing of the actual world that it is so and so..." (Soames 2002: 327, n. 43) – it becomes obvious that that characterization of the content objects of the *de re* propositional attitudes, i.e. them being about the actual world objects and their properties, does not differentiate the objects of proper names from the objects of the actualized definite descriptions because both are about actual individuals. So, *pace* Soames, from the sentence, "A believes that the actual shortest spy is a spy" (assuming spies actually exist) it follows that there is a spy that A believes to be the shortest, just as it follows from the sentence "A believes that NN is the shortest spy". To know who the shortest spy is, equivalently who NN is, is not required; and being *de re* is not precluded from the definite descriptions either: the causal contact must be relevantly informational, i.e. not just whatever fleeting physical relation with the referent of the name and the description.

The fifth argument is finally Soames' own and we need to set out its premises at length (Soames 2002: 43-4):

P1. It is possible to believe that Aristotle was a philosopher without believing anything about the actual world A_w – that is, about the way the universe really is (the property it really instantiates). In particular, there are worlds w^* in which agents believe that Aristotle was a philosopher,

without believing of A_w that anything was F in it, and hence without believing of A_w that the unique thing that was F in it was a philosopher.

P2. Necessarily, one believes that the actual F was a philosopher iff one believes of the actual world, A_w , that the unique thing that was F in it was a philosopher.

C1. It is not the case that, necessarily, one believes that Aristotle was a philosopher iff one believes that the actual F was a philosopher.

P3. If the content of *Aristotle*, as used in a context C, were identical with the content of *the actual F*, as used in C, then (i) the contents of (propositions expressed by) *Aristotle was G* and *The actual F was G* in C would be the same; (ii) the propositions expressed by α *believes that Aristotle was G* and α *believes that the actual F was G*, in C, would be necessarily equivalent; and (iii) C1 would be false.

C2. The content of *Aristotle*, as used in a context C, is not the same as the content of *the actual F* as used in that context.

Soames claims that each premise is true. But it seems to me that that is not so: P1 seems to be false – or alternatively irrelevant to the issue of rigidity. Of P1 Soames says that “Surely it is a datum that agents could have believed that Aristotle was a philosopher even if things had been quite different from the way they in fact are. Must these agents also have had beliefs about the actual world...Presumably, in some merely possible world (world-state) the agents there have no direct acquaintance, or epistemic contact, with this world (world-state) that I am now calling ‘actual’; nor, in many cases, will they possess any uniquely identifying descriptions of it. As a result, often there will be no way for them to form beliefs about the actual world.” (Soames 2002: 44) But what is Soames’ argument then supposed to be? What is it supposed to be showing? If there is no contact whatsoever with the actual world (this world of ours here now), what do those otherworldly agents believe when they believe that *Aristotle* was a philosopher? For the crux behind the rigidity is that “Aristotle” refers rigidly to the *actual* Aristotle (in our world here now as well as in every other possible world where *that* Aristotle exists). Even if the syntactical formulation of their belief is the same as ours, there just is no way to make sense what those otherworldly agents believe if they are not related to the actual Aristotle in any way (short of accepting interworld magical theories of reference). And as Soames himself states they do not believe anything about the actual (our here) world. If they did, the expression “Aristotle” describing their belief would also be about the actual world (given the rigidity of proper names). Hence Soames’ premise P1 is not sustainable. Or it is irrelevant to the issue of rigidity: if proper names are not rigid, then the actualized descriptions view is not even needed. But as I argued above it also manifests the referential force, the counterfactual scope, implicit in the ordinary referential uses of the definite descriptions. So those descriptions would *be* rigid in that use. Either way, Soames’ argument fails. If proper names are rigid then the actualized descriptions are equivalently rigid too. So either the otherworldly agents believe that the actual Aristotle was a philosopher or we do not know what they believe when we express their belief by the statement “Aristotle was a philosopher”. Consequently Soames’ premise involves an incoherence with respect to rigidity.

It also seems that Soames have misunderstood the role of the “actuality operator”. For he seems to assume that the only acceptable reading of “the actual F” always restricts it to the world (the context) in which it is used. From this it follows rather trivially that whatever was the actual F in that world, and who was a G (as “our” actual Aristotle was a G, by hypothesis), that “actual F” and “Aristotle” need not be co-referential. But here Soames has assumed that “Aristotle” is rigid and the otherworldly description “the actual F” is not rigid in the same way. But the whole point of the rigidity issue is that all referential uses are anchored *in our world here*, those of proper names as well as those of the *actualized* definite descriptions. Hence it is invalid to read “the actual F” as anchored on the other world to begin with. The actualized descriptions view is only about the descriptions of the objects in our actual world. It is beside the point that they are evaluated in (or at) other possible worlds, or counterfactual situations. Kripke has been explicit in emphasizing that it does not matter what the expressions of the language spoken by the agents in the counterfactual situation refer to or mean *there*; it only matters to what *our* expressions refer as evaluated in that counterfactual situation. (See also (Stanley 1997: 555-6).) In other words, why should we care what Soames’ otherworlders believe and say from their perspective? All in all, Soames’ conclusion C2 does not follow.^{11,12}

The sixth argument by Soames is the following. Soames acknowledges that actualizing a definite description is equivalent to the application of David Kaplan’s “Dthat” operator (Soames 2002: 49-50). However he argues that this move does not help because Dthat obliterates the descriptive content of the description it is attached to. From this it is supposed to follow that the descriptions could not be called for help any more after the fixing of Dthat to them because then, when someone believes that NN_1 is F but does not believe that NN_2 is F (and “ NN_1 ” is co-referential to “ NN_2 ”) that could not be explained for the reason that the descriptive content is supposed to be there no more because of the application of “Dthat”. But this seems again a misunderstanding, or rather, perhaps, a misapplication, of the “Dthat” operator. “Dthat” just rigidifies the description in the way that its uses refer to the actual individual the description applies to. (It need not even apply to that individual truly; it suffices that it is believed to apply.) Just as a proper name is the vehicle that ties its uses to its referent rigidly, so is a “Dthated” description the vehicle that ties it rigidly to its actual referent. The description is not re-assigned a referent in any counterfactual situation after it has been actualized in our world (or with respect to our world). It is not denied by the actualized description theorists that there could be in some other possible world, say, the most famous pupil of Plato who is a philosopher, but who is not the actual Aristotle. Still the actualized description “the most famous pupil of Plato” picks up the actual Aristotle in every possible world (where he exists). The application of “Dthat” to the definite description just does not, as Soames seems to think, take away the descriptive content related to the proper name and make it, in effect, a mere actualized demonstrative. The descriptive content and the demonstrative factors work together with equal force in “Dthated” statements. And Kaplan himself has pointed out that “Dthat” can be eliminated in favour of the definite descriptions with the “actuality” and “now” operators (Kaplan 1989: 579, n. 29). So I conclude that Soames’ sixth argument fails also.

Conclusion

So all in all, if my arguments are correct, the actualized descriptions view remains unharmed by the Kripkean modal argument as well as by Soames' argument from the mesh of the modal and propositional attitude contexts. Kripke's argument fails because it focuses only on the predicative reading of the definite descriptions. Soames' argument fails because it invokes assumptions that violate the very point of the (Kripkean) notion of rigid designation: other worlds are strictly speaking irrelevant to the references of the rigidly referring expressions. In particular it is beside the point to put any emphasis on every possible world being actual with respect to itself and then explicate "the actual such-and-such" statements with respect to that world. In rigidity issues only our actual world counts with respect to both the modal and the propositional attitude contexts. That could not be otherwise because they *both* are intensional contexts, so adding the propositional attitudes to the modal ones adds nothing of substance.¹³

As to the relevance of my argumentation to the Fregean sense theory the connection is straightforward. If the definite descriptions express the senses of proper names related to them, when the descriptions used as actualized have attached to them - through their referential-actualized use - the referential force equivalent to rigidity, the senses they express have that force too. So the modal argument does not reach the Fregean theory either. This verdict is not changed when the inarticulatable elements of the senses are taken into account, for they do not effect a sudden "jump" in the intended referential scope from our actual world to some other possible world (considered as another actual world). The remaining difference between proper names and the definite descriptions manifested in the pairs of statements like 1) and 2) above was explained earlier, and will be accounted for a little more exhaustively in the next chapter. But let me take it up here in outlines. The differential modal impression is the result of not keeping separate the referential-actualized and the predicative readings of the descriptions. The sentence "the most famous ancient philosopher might not have been the most famous ancient philosopher" is contradictory when read predicatively throughout, i.e. when the focus is exclusively on the descriptive content of the description. But when the description is read in the mode of referential-actualized, and predicative respectively, it is a true statement (barring considerations from strict determinism) for Aristotle might have gone in for *pankratíon* instead of philosophy. The differential effect comes from two sources. First is the statistical fact that the definite descriptions are also frequently used predicatively ("NN is the such-and-such"); and second one is that because the descriptions express the information the speakers possess about the referents of proper names (and so express their senses and meanings), this predicative role of the descriptions tends to influence the statements in which they are used. Consequently because proper names tend to activate (many of) the pieces of information (=the elements or parts of the senses of those names) the descriptions express, the predicative reading "forces its way" into the reading of the statements. But when a definite description is used referentially and actualized that description is so to say chosen from amongst the other related descriptions (the whole of the information related to the proper name in question). It is this interrelatedness of the descriptions as the sense-expressers that makes it hard to suspend the predicative reading from being taken as primary.

Notes

1. Is it impossible to hear (to “intuit”) the predicative reading of “Aristotle could not but have been Aristotle”? I think not, for I can comprehend what that reading says – with effort, though. So I claim that it is possible and the effort one needs only reflects the degree of seldomness of the predicative use of proper names, but no in-principle feature of them.
2. There are places in Russell’s writings that can be interpreted according to the actualized descriptions view. See for example (Russell 1956: 52-3). As Russell explains the “primary occurrence” of a definite description is intended by the speaker when one and only one X is intended and to that X something is attributed. Thus George IV wished to know who wrote *Waverley*. The “secondary occurrence” reads, instead, that George IV wished to know whether one and only one man wrote *Waverley* (and if so whether that was Scott). Let me quote Russell in full: “To return to George IV and *Waverley*, when we say, ‘George IV wished to know whether Scott was the author of *Waverley*’, we normally mean ‘George IV wished to know whether one and only one man wrote *Waverley* and Scott was that man’; but we *may* also mean: ‘one and only one man wrote *Waverley*, and George IV wished to know whether Scott was that man’. In the latter ‘The author of *Waverley*’ has a *primary* occurrence; in the former, a *secondary*. The latter might be expressed by ‘George IV wished to know, concerning the man who in fact wrote *Waverley*, whether he was Scott.’” I claim that as far as the ordinary referential uses of the definite descriptions (Russell’s “denoting phrases”) are concerned the primary occurrence is equivalent to the actualized descriptions. There seems to be no doubt about that equivalence in their referential forces: Russell’s use of the locution “who in fact wrote” indicates this when “in fact” is semantically equivalent to “actually”. There seems to be no reasonable alternative interpretation whatsoever with which to try to counter the natural reading that “in fact wrote” is intended as saying (or paraphrasing, if one likes the formal mode better) that the referent is the one who actually wrote *Waverley*. The secondary occurrence, in contrast, focuses on the descriptive-cum-predicative content of “the author of *Waverley*”. As Russell says the secondary reading amounts to finding out who wrote the piece, if one man wrote it. (Though it may be that even here the referential intention has the force of “in fact”, which only strengthens my case.) – Be that as it may with Russell, Strawson’s theory can be easily interpreted as a version of the actualized descriptions theory. For Strawson maintains that there is a presupposition of existence of the object referred to (Strawson 1971b). When the purported object does not exist, the statement the part of which the referential term is does not have truth-value. Because of that I take it to be natural that in the latter type of cases the referential expressions – descriptions in our case – are not to be evaluated with respect to any other counterfactual situation, either. The existence presupposition can be taken to amount to the descriptions having the semantical force of the actualized descriptions; Strawson just eschews them any counterfactual application. – as an afterthought taking into account that Russell did not care about the modal contexts (let alone possible worlds) his theory’s clause “there is at least and at most one individual such-and-such” is evidently restricted to our actual world only. This supports indirectly the actualized descriptions view. But even if not, as far as I can see there is no semantical reason that would prevent a construction of a version of Russell’s theory such that would possess the force of the actualized descriptions (and on account of that be modally extended).

3. Davies & Humberstone assume that proper names, in contrast to the definite descriptions, need no world relativization (Davies & Humberstone 1980: 12). But as we saw that is only a manifestation of the widespread misconception. Plantinga's argument invokes essences of individuals. These constructions I could not accept (if they are meant to be anything stronger notions than akin to the individual *concepts*). But his argument remains valid. Notice that I am not saying that the wide scope reading is the same as the actualized descriptions view. They are not because a description can have a wide reading in a possible world without that world being the actual world, i.e. our world. (Though I am not sure whether or not that type of situation allows *us* to be more specific at all about the other world. The reason is that we could only talk about it by using *our* language, i.e. the references its expressions have, and these need not be the same at all in that other world.) –We saw in section 4.2.4 that proper names have also predicative use (“He is a regular Koba”), albeit infrequent. So the argumentatively harmful way Kripke presents the modal argument becomes even more aggravated: he concentrates only on the referential function of proper names and only on the predicative reading (of the referential function) of the definite descriptions. Consequently it is no wonder that he is able to make the modal argument seem as plausible as it has seemed to so many. Instead he should have kept the referential reading stable through the whole argument; but in that case he would have realized that that is equivalent to the actualized-rigidified descriptions, hence the modal argument becoming void. (Note that Kripke's, and the other direct-causalists' point *could not* be that proper names *are* rigid as such. There could not be predicative uses of them if they were only rigidly referential as syntactico-semantic devices. This is the same reason why Recanati's argument for the type-referentiality of the directly referential expressions was found wanting in section 4.2.4.) Shortly, 1) “Aristotle might not have been Aristotle” can be taken as false; the statement 2) “The most famous ancient philosopher might not have been the most famous ancient philosopher” has combinations of readings: i) predicative-predicative, ii) predicative-referential/actualized, iii) referential/actualized-predicative and iv) referential/actualized-referential/actualized. According to i) 2) is false in the same way as 1); reading ii) says that some individual, whoever that is, who is the most famous ancient philosopher might not have been Aristotle, and that is true (because there could be such a world); iii) reading is equivalent to saying that Aristotle might not have been the most famous ancient philosopher which is true also; and finally iv) says that Aristotle might not have been Aristotle, which is as false as 1), for it says what 1) says. Now it is easy to see that the modal argument fails because the only relevant reading, with respect to that argument is iii) actualized-predicative, but that reading is not forced upon the description theorist: he can choose the reading iv) that is equivalent to the rigidity claim the statement 1) involves. So only by (inadvertently, I hope) insisting on the predicative reading of the definite descriptions can Kripke claim that – even after the “fixed understanding” that the proper name in question refers to an already individuated and identified object – “rigidity requirement is violated” in Russell's theory with respect to the definite descriptions (Kripke 1980: 9). But with respect to the actualized-referential reading this claim has no bite.
4. The core tenet of the thesis of proper names as rigid designators is that they refer to the same actual world individual in every possible world (counterfactual situation) where that individual exists. Why, then, are not the definite descriptions allowed that? For surely there are different individuals named “NN” in those different possible

worlds, just as there are different individuals who are “such-and-such” in them. Only if the definite descriptions are read as predicative to begin with, and proper names as (only) referential-rigid (and the relativization of use is taken as unproblematic), could the former be put in use in the modal argument. But that reveals the modal argument to be *non sequitur*: the definite descriptions can be used referentially and *with* the force of rigidity due to their tacit actualization-through-use. Kripke’s formulation of the modal argument is defective also in that it uses only *one* description, when Kripke’s attack was directed against the *cluster* version. So the proper formulation should include *all* that is known and related to the named individual (for example, Aristotle) at the same instant (and not all those descriptions separately). But then it begins to look that the content of the descriptive part approaches the rigid referential force the proper name has. (Which is, of course, evident: according to the description theory the proper name contains implicitly those descriptions.) Hence the reading that “Aristotle might not have been D_1, D_2, D_3, \dots ” (D_i s being the relevant descriptions) becomes as false as “Aristotle might not have been Aristotle”. (At least that is what my “intuition” strongly tells me.) (As we saw Salmon’s formulation takes the multiplicity of the descriptions into account – but that does not change the verdict. And notice that I am *not* saying that the definite descriptions as such, i.e. as descriptive devices with the focus only on their explicitly expressed content, are always “locked” to a particular actual individual in their actualized uses, in contrast to other actual individuals. But, then, neither are proper names always so: they have to be disambiguated (relativized in use, in context). And that may require non-conceptual and contextual factors that are not linguistically articulatable.)

5. Synonymies and analyticities between proper names and “their” descriptions is an independent issue from the referential-predicative readings: synonymical description can be used both referentially and predicatively. (So for example Soames is both irrelevant and in error when he states that actualized-rigidified descriptions are synonymous with proper names (they relate to) (Soames 2002: 40).)
6. In case you need 4 points clue: In 1926 she disappeared inexplicably for 11 days. Many thought that was a publicity stunt. (3 points: Her play “Mousetrap” has been on stage in London since 1952 – more than 21000 shows. 2 points: The famous auction established in 1766 has the same name. And 1 point: Among the characters she created are Hercule Poirot and Jane Marple.)
7. So for example Stephen Schiffer is partly correct and partly wrong when he states that the actuality operator is “pretty sophisticated” and that “one might reasonably doubt that it belongs to the conceptual repertoire of young children and other non-philosophers.” (Schiffer 2005: 1175) He is correct in that the operator is not explicit (except in pointed semantico-logical studies, but young children rarely engage in those). He is wrong, however, because being part of the repertoire need not mean being explicable or even consciously used. (That is why the rigidity force is tacit with the actualized definite descriptions.) Again, by the way, how young a child must be to not manage the actuality-rigidifying force? It seems irrelevant to the issue when any speaker tacitly assumes the actuality operation. Schiffer also mentions Soames-type counterarguments to the actualized description view, but we will see shortly that they fail.

8. Donnellan's distinction is partly based on this misconception. For how otherwise to explain that he contrasts the attributive uses of the definite descriptions with such referential uses that work only in demonstrative situations. That is, if he had realized that the definite descriptions *are* used referentially with actualized force, the attributive use would not have seemed so important to him. Or so I presume. Of course Donnellan's focus on Russell's theory of definite descriptions played the other part in his argumentation, but as I mentioned earlier, even Russell's theory could be interpreted as supporting the actualized descriptions view.
9. Soames has recently argued against the so-called "two-dimensional semantics" (Soames 2005). That issue is not relevant to my argument here for, as Soames himself acknowledges, the actualized descriptions view is distinct from the two-dimensional semantics. They are related because that semantics also uses actualized descriptions as rigidified referential devices. But Soames only repeats – (Soames 2005: 303–4) – his earlier arguments against the actualized descriptions view, arguments that I counter in this section, so we need not pay a further attention to his recent book. (Moreover, two-dimensionalism is motivated mainly by the epistemological-modal scruples – contingent *a priori* and necessary *a posteriori* – which are altogether different issues than the reference of proper names. The latter is primary, the epistemic-modal are mere corollaries the chips of which I can let fall where they may. But to the extent that I admit that there are rigid designators I can have it the Kripkean way, i.e. that there are necessary *a posteriori* truths. But if there are not, the notion of rigid designator (proper names and actualized descriptions – and "Dthated" terms, see the end of section 5.3) remains a sound notion as explained in the next chapter (and at section 4.5). – Soames' arguments are defective in that he still clutches to the mistaken view that the descriptions must not contain any uneliminable object-dependent elements (like other proper names), and that the actualized descriptions must be synonymous with proper names (Soames 2005: 268).
10. As noted Soames claims that the actualized descriptions view is after preserving synonymies between proper names and the actualized-rigidified descriptions (Soames 2002: 39, 40). But this is a misconception, and I think that merely pointing this out one more time by now suffices. Of course it is most unfortunate that the synonymy demand for the descriptions (and for the Fregean senses) is still so entrenched in the direct-causal camp. However I will deal with Soames' arguments as if they were independent on the synonymy demand.
11. Soames claims that the propositions *the actual F was G* and *the F was G* are different, say something different (Soames 2002: 45). That is true but here he equivocates the referential actualized use with the predicative use of the descriptions. And another point, related to the fifth argument: Soames claims that English belief ascriptions are understood as being true at arbitrary worlds only if the agents' beliefs in those worlds are about the individual in (our here) actual world, "...even though the beliefs themselves are not about [the actual world]" (Soames 2002: 49). How could they not be about the actual (our here) world if they are at least partly about the actual individual in our actual world? Is not the case symmetric, i.e. do not we have beliefs about individuals in other worlds but not beliefs about those worlds? But how could that be, unless talking about beliefs about particular individuals can only be made sense by its being talk about actual world individuals? But then actualized descriptions *are* as much in the business as proper names. So what distinction Soames is making use of

when he differentiates beliefs about individuals from beliefs about worlds? (It is hard for me to comprehend Soames' argumentation in this place for it seems so elliptical that I am not sure whether it is he who is confused or I.) This is even more confusing on Soames' part because he also says that his argument "...is based on the elementary observation that not only individuals in the actual world, A_w , but also inhabitants of other worlds, share many of my beliefs." (Soames 2002: 43) For if they *share* those beliefs and those beliefs are about the actual world (as Soames' beliefs in this world about this world are, say, his descriptive beliefs about the actual Aristotle), the otherworldly agents have beliefs about the (our here) actual world.

12. By the same token of argument I could not accept Recanati's argument that descriptions are non-rigid when they fall outside the scope of modal operators (Recanati 1988: 106; 1993: 8-10). According to his argument, if every world is actual to itself, it is possible that there is no unique individual in all other (relevant, accessible) possible worlds that the description specifies. But this interpretation of actuality affects the bearers of proper names as well. If "F. Recanati" in some actually taken world, not *our* actual world, is a rigid designator, the referent of it in all other (relevant, accessible) possible worlds is *that* FR, and there is no guarantee that *that* FR is the FR in our world. (Notice by the way that "the" does not just restrict definite descriptions to unique individuals, but also proper names: this indicates the tacit application of disambiguating proper names. And related to that observation proper names can be actualized also: that we have been doing all along in this chapter when we have been talking about "the actual Aristotle".) So "FR" is not a rigid designator in Recanati's argument because it severs the notion of rigidity from being anchored only in our actual world. Another way to put this is that Recanati's argument ignores the relativization of the use of proper names, i.e. the individuation of their referents before any rigidity claims could be made - which we saw is the common mistake by the direct-causal theorists also. Recanati is assuming that it is adequate to assume that the actuality operator can be restricted to apply to one world other than our actual world at a time, and this seems to be the same mistake as Soames makes.
13. Their effective equivalence with respect to semantic matters have been repeatedly argued for by Jaakko Hintikka; see for example (Hintikka 1969; 1975).

6. FROM NEO-FREGEANISM TO NEURONATURALISTIC FREGEANISM

Problems with neo-Fregeanism

I have been arguing that the notion of the Fregean sense of proper names is a sound notion with respect to all the tasks it was postulated to explain. That this is so is what the neo-Fregeans have *assumed* as the starting point of their studies, so compared to them I am on a firmer ground. Moreover it seems to me that the conceptions of the senses the neo-Fregeans have invoked are many and varied, without any evidently undisputably common property. In particular we still lack adequate answers to two questions: 1) what specifically the senses of proper names are?, and 2) what is to grasp such senses? We have inherited these problems from Frege because he did not say anything much non-metaphorical about them. Basically the grasping of a sense is a psychological act that allows a thinker to confer something of cognitive relevance to one's linguistic expressions, *a fortiori* to proper names.

In the neo-Fregean camp Michael Luntley has recently voiced the need for some model which would explain the *recognition* of a person as a cognitive achievement, and which would as well be explicative of the senses of proper names (Luntley 1999: 253). This is in accord with the views of Gareth Evans and Michael Dummett that recognition of individuals manifests an important function of the senses. More particularly Luntley states, in relation to the partial inarticulatability of the senses - i.e. the unattainability of fully linguistically articulated individuation of the senses, which is currently acknowledged to be part of the neo-Fregean approaches - that

"We can distinguish between the theory of sense and a theory of senses. The latter is a theory capable of specifying, for any given linguistic unit, the sense of that unit...In contrast, the theory of sense has no such aspirations. The theory of sense specifies the kind of thing sense is: its fundamental role within our self-conception as creatures with intentionality, what kind of information sense is, and how this will differ systematically for different categories of expressions...The theory of sense must give some account of how we could understand the claim that each singular term has an objective package of information associated with it that constituted its sense. *It does not follow from this that it should be always possible to state what that package is...*" (Luntley 1999: 231-2; my emphasis)

I endorse this approach and try to give such an account of the general nature of the senses of proper names in the next two chapters. As we will see, indeed, the senses of proper names can not be always articulated without any remainder in linguistic terms, in particular by descriptions.

But as said, in general the neo-Fregean studies, valuable though they are, have been pursued in a rather "phenomenological" way. Usually the neo-Fregeans have been content with quite general pronouncements like the ones that a sense of an expression is what the thinker knows (or has to know) in order to be able to determine the referent of that expression. The more penetrating studies have invoked folk psychological concepts and secretarial analogies like information about the referents of proper names

contained in "files", "dossiers" (Evans 1982; Forbes 1989, 1990; Recanati 1993), and "clusters" (Kvart 1989) (though Kvart's clusters consist of linguistic expressions only, so his account is rather a mixture of the description theory and Fregean view, but the former carrying the most weight). These analogies are catchy, but the more specific, and dynamical, relations between proper names and the information related to them still remain quite unexplored. However, the positive point of the neo-Fregean secretarial analogies is that they do invoke information, thus indicating that the (future neurocognitive) accounts in which they are (to be) embedded are not exhaustively description-theoretical ones, thus disarming the persistent but mistaken view most of the direct-causal theorists still seem to hold.

So it can be claimed that to some extent the secretarial analogies do manifest progress within the neo-Fregean camp - though for instance Evans' original conception of "Ideas" of referents remains quite unarticulated despite the fact that he discusses extensively recognition and memory-based thoughts. For instance he does not penetrate into the specific cognitive dynamics of those mechanisms, but uses examples at the "phenomenological", molar-behavioral level together with folkish guides as to what seems to be the case. ("we would say that" and "we would not say that"). The largest gap this approach leads to is Evans' total rejection as being relevant what goes on in the brain in the recognition- and memory-based processes. These processes he takes to be very different from what the *subject* does. This gap only grows because Evans argues in part against the description-theoretic approach and, though correctly, emphasizes the role of information in the referential individuation of the referents in thoughts. This distinction explains in large part Evans' use of the phenomenological level concepts. For if it is the case that what only matters is what *persons* do, and are capable of doing, one is almost conceptually hindered from taking the next step toward the specific cognitive models - let alone the neuroscientific ones. But because of that hindrance the importance of the informational focus is bound to remain limited. To a comparable extent folkish studies will remain insufficient explanations of the specific mechanisms of the reference of proper names, as well as reference in general.

Another argument for the neural-cum-cognitive step presents the neo-Fregeans in a dilemma. *Information is information*; i.e. information at the phenomenological level is what the information processing paradigm in the cognitive sciences focuses on, so the at first progressive informational step by the neo-Fregeans is in danger of turning to a regressive one. Or if information is *not* information, i.e. if the phenomenological level of information is different from the neural-cum-cognitive information, the neo-Fregeans owe us still an argument that would show the *primacy* of the first type of information. But from my modern Fregean perspective it could only be the study of the informational structures and dynamics within the subject that enables us to make further progress. The main reason for the gap between the person and his brain seems to me to be that Evans indeed made a categorical distinction between the acts the subject performs and the informational processes in his brain.¹ But such a distinction is an outdated one in the current theoretical climate; in the philosophy of language and mind there are no obstacles to the explanatory top-down approach in which we start with the person and by descending from that level down, step by step, we reach the specific neural systems executing the perceptual and cognitive functions that make up what persons are. To ignore the person-brain divide is one of the main motivations of my approach in this work: as I see it only the neurocognitive approach can help us

forward. That this approach is to some extent top-down but more bottom-up is all to the good explanationwise.

The observation that in addition to the cognitive roles the senses have epistemic roles as well was reinforced by Michael Dummett (1973, 1981). He takes the sense of a proper name to be what a person knows when he understands that expression, or its contribution to the truth-value of the sentence, or an utterance, of which it is a part. Dummett also stated that

"The sense of an expression is the mode of presentation of the referent: in saying what the referent is, we have to choose a particular way of saying this...we might here borrow a famous pair of terms from the *Tractatus*, and say that, for Frege, we *say* what the referent of a word is, and thereby *show* what its sense is." (Dummett 1973: 227)

And Evans says, by way of an explication, that

"The closest we shall get to a systematic theory of sense is a systematic theory of semantic value which, however, identifies the semantic values of expressions, including whole sentences, in the way in which the competent speakers of the language identify them." (Evans 1982: 26)²

For example with respect to the clauses 1) "the semantic value of 'Afla' = Afla", and 2) "the semantic value of 'Afla' = Ateb", only the first shows the sense of "Afla". It should be pointed out that Dummett's view requires only that it is sufficient for any term *t* and a person *p* to have a sense *s* by which he understands *t* (Dummett 1973), and not, as Evans did have it, that for any term *t* there is a sense *s* such that for any person *p* to understand *t* is to attach *s* to it (Evans 1982: 16, n. 14). That is, it is not required by Dummett, as it is required by Evans, that there is one particular sense shared by every speaker competent with the expression *t*. In particular Evans states that sense is the particular way one thinks of the referent, and to understand what someone says is to think of that referent in the same way. I side with Dummett on this issue because his view is Frege's view, according to which different persons may have different senses related to a particular term *t*. (See note 3 in chapter 4.) The evidence for this is Frege's Aristotle note, again. But on the whole I think that the way both Evans and Dummett understand that saying what the referent is, is by showing the sense of the expression, is not adequate from the properly Fregean perspective. For what difference is there supposed to be between "Afla = Afla" and "Afla = Ateb" such that only the first statement shows the sense but not the second one? Suppose that the two geographers have the same referent determining experiences in Frege's original example but only the names they relate to the mountain differ from each other. In that case the identity statements also show the same senses (excepting metalinguistic ingredients). It seems to me that the correct way to understand the identity statements is rather the following. "Afla" and "Ateb" in the right side of "=" bring always with them the associated senses of an individual speaker. But that takes place below the linguistic surface level, and therefore what is shown is not really *shown* at all. In most cases when different proper names are explicitly given they can be taken to *indicate* that different senses are involved but not that the senses are shown, as it seems that Evans' and Dummett's views require.

So to talk about showing the sense of a proper name is too coarse, for that does not really enable us in finding out at the linguistic level what the senses of proper names are the speakers grasp. The plain talk of showing leaves it totally open how the sense is shown when this is meant to be contrasted to the possibility of stating the sense directly, albeit with linguistic means. This relates to the problem of grasping a sense of a proper name. Consequently I propose that showing the sense reduces empirically, not conceptually, to a manner of grasping a sense.^{3,4} And what that is, in modern cognitive terms, will be explained piecemeal further on.

The general line of my argumentation is then the following. The information-based thoughts that Evans focused on is an important step forward in the neo-Fregean studies.⁵ I would even claim that this notion manifests a very similar overall account to mine in that it also plays down the role of causal links *qua* causal (even if inadvertently to some extent). Only information counts; the causal links are at most its carriers. (That is, when the links are conceptualized folkishly to be distinct from the information – which may not be, perhaps not even metaphysically, possible at all, as I argued in section 4.2.5.) But there is still a discrepancy between the senses of proper names and the demonstratives and indexicals in Evans' account. He discussed proper names in his 1982 book, but the discussion remained on the intuitive level of what we would say with respect to, for instance, when a proper name has definitely changed its referent. Evans' analogy of the producers and the consumers of a proper name does not really add any explanatory value; it is only a cogent description of the social dynamics of reference change. Fortunately there is a later development within the neo-Fregean camp that focuses also on the senses of proper names.

Recanati's account

The notion of a sense of a proper name fluctuates quite much in the earlier neo-Fregean pronouncements and tends to remain rather general. But fortunately there is a more recent and more developed account of the neo-Fregean senses of proper names. We saw earlier that Recanati talked about the "architectural trick" of putting information about the bearer of a proper name into a dossier. We can drop the secretarial analogies from now on. That is because we can get much more literal: we do not "put" information to the "dossiers". Information goes into the neural memory stores in the cortex along the neural links that has been formed, or that are formed at the moment when one acquires a new proper name and information about its referent. Talking about putting information in a dossier makes the whole process turn out to look as something that we do in full awareness down to the last piece. But that is not how it goes neurally, and not even cognitively in most cases. To venture a bit to the neural territory of the next chapter, the brain gets the information from which it selects pieces in accordance of what it has already acquired, and consolidates the information by transferring it from hippocampus to the cortical networks. This is essentially the "architectural trick". Recanati's account is also one-sided, for it needs to be explained how the informational traffic goes from the other direction, from the cognitive memory information (the "dossiers") to the linguistic expression of a proper name. But that will be explained also in due course.

Still the general features of Recanati's account of the senses of proper names are valuable, for they clearly make a step forward. Recanati makes it clear that "a dossier" can be initiated by a referent-fixing description (when in effect the proper name is a

"descriptive name"). But when afterwards new information is acquired by a thinker, the original description "gets subordinated to the dossier and loses its privileged status: the name becomes associated with the dossier rather than with the description which initiated it." (Recanati 1993: 110) This leads to the situation in which the referent of the proper name is thought of "non-descriptively" (possibly without the thinker ever having been perceptually acquainted with the referent). In short, the non-descriptive mode of thinking of the referent is the result of "sheer descriptive multiplicity". This mode of thinking ("*de re* representation") about the referent is such that the thought characterizes the referent itself independently of it satisfying the concepts that are entertained by the speaker in thinking of that referent. That mode is the "mental counterpart" of the linguistic feature REF that characterizes the directly referential singular terms.

The non-descriptive mode of thinking about the referent through the descriptive multiplicity is, however, left much too unarticulated by Recanati. It even seems to have something mysterious about it in cognitive terms. For plainly, if there are nothing but descriptions in the dossier, how their sheer *amount* could give rise to a *qualitatively* distinct mode of apprehending the referent? When the descriptions accumulate, at what point does the mode of presentation change into non-descriptional? And more pressingly, *how* could a descriptional mode turn into non-descriptional mode? I have not found any answers in Recanati's discussion. But in a way I do not expect to find them, for it seems to me that Recanati has confused cognitive matters (his psychological modes of presentation) to epistemic ones. Accordingly I propose that instead of the descriptional mode changing into the non-descriptional, what happens when one acquires a *de re* mode of thought about an individual is that the accumulation of the descriptions increases their *mutual cohesiveness*. That is, the descriptions become to express such pieces of information that enforce one another as being about the same individual. That is natural: it happens (almost every time) one acquires new descriptions – new information - about an individual, and even in those cases when one knows that those descriptions are about the very same individual. In other words one's sense of the referent (of the proper name in question) becomes *epistemically reliable* in its overall informational coherence to such an extent that it is taken to characterize a unique individual. We see again that semantic, *a fortiori* referential, and epistemic considerations mesh; but then that tenet lies at the core of our Fregean approach also.

There is another level of the senses of proper names that contributes to the epistemic cohesion, hence to the Recanatian *de re* mode of thinking about the referents. It consists of the informational relations between i) the particular senses and ii) the senses about other things, like events in the world, i.e. the world knowledge the speakers possess both individually and collectively through communication. In the modern Fregeanism the senses of proper names are not separate entities but only relatively separable because informationally interrelated. This second level of the senses supports evidently the cognitive and epistemic nature of reference. Consequently we also see that this epistemic base supports internalism over externalism. (See sections from 4.3.1 to 4.3.3.) The senses are related both to the external world and to one another, but the former connections are not so strong as the externalists believe. That is because the latter connections constrain the former, i.e. the connections within the internal representations as well as to the external world are also informational. And because this information, the senses, are our only means by which

we know anything about the external world, nothing from outside the individual representational system made up of the senses could mediate cognitive, hence semantic, relations between them and the external world. In short, Fregean internalism is the only viable view in semantics and reference research.

Still it needs to be remarked that strictly speaking the dossiers do not contain linguistic descriptions, *pace* Recanati and Kvat. The descriptions are the linguistic items one produces from the cognitive information the dossier contain. Recanati does appreciate however this point for he says that the descriptions express concepts and that we could not think about objects without a conceptual mediation (Recanati 1993: 130, 186). But it seems that the “dossiers” contain also perceptual information about the referents. To see this, think about exemplars and typical instances of categories and concepts.⁶ Recanati also realizes that one acquires information about the referent via perceptual encounters with it. Nevertheless in his account that seems only to result in the linguistically articulatable, and representable, pieces of conceptualized information entering the dossiers. In light of this Recanati’s notion of “encyclopedia entry” needs modification. Consequently it is better to take the notion of the dossier as a cross-informational and a functional construct. The same goes with the encyclopedia entries informationally explicative of the dossiers about individual referents (Recanati 1993: 184). In other words the encyclopedia entry contains one’s knowledge and beliefs about the referent of the expression, including the “pointer” by which one gets access to the entry at a time. It is also to be understood that there may be more than one proper name linked to one and the same entry; in that case the two names have the same sense (excepting of course those portions of the sense which have to do with the fact that the names differ).

The relevance of proper names as pointers, with respect to the encyclopedia entries, is that with their help the speaker both initiates and accesses the entries. (For similar accounts see (Forbes 1989) and (Kvat 1989).) But the entries can also be initiated and accessed by descriptions (as well as with indexicals, demonstratives and pronouns). The basic elements of the initiation was described above with the acquiring of the original description: it causes a “peg” for “clustering” further information about the object that the very information is about (Recanati 1993: 183). The access in its turn goes via any information the entry contains, as when one hears something about the referent that is already known to him. And obviously one can acquire new information in that occasion.

Let me try to make the notion of encyclopedia entry, as the secretarial analogy of the Fregean sense, more clear by an example. For this role I offer portions of my sense of “Stalin”.

- “Stalin” (the main proper name pointer); “Josif Vissarionovich Dzugashvili” (birth name), “Soso”, “Koba” (other proper name pointers)
- born officially 21st December 1879, really 8th December 1878, at Gori, Georgia
- met Lenin first at Tampere, Finland, December 1905
- Commissar of Nationalities in 1917
- second wife Nadezda Alliluev
- Secretary of the Communist Party of SU
- vozhd* of the Communist Party and Soviet Union since 1929

- started the first *pyatiletki* in 1929 (5 years plan)
- worked closely with Molotov
- started the massive campaigns to industrialize and dekulakize SU in the beginning of 1930s
- let few million Ukrainians die of hunger
- originated the purges of the higher party, economic and Red Army echelons in 1936-1938 in the show (and secret) trials
- negotiated with the Finnish envoys in October and November 1939; as the result started the Soviet-Finnish War ("The Winter War") 30.11.1939
- deliberately on purpose let Germany attack SU in 1941
- supreme commander (at *Stavka*) in the "Great Patriotic War"; later assumed the title Generalissimus
- died at 5th March 1953

This encyclopedia entry lists some main properties of Stalin and events in his life. All of them are not separately sufficient for unique individuation of him, however. But that is a common feature of the senses: in many cases they contain pieces of information that alone would not suffice for identification of the referent of the proper name they are the sense elements of; but together they are sufficient. That can be achieved in two ways. Either the pieces of information together form uniquely individuating sense or there are some elements that suffice for that alone. With respect to my sense of Stalin such are, for instance, that Nadezda Alliluev was his second wife, and he being the supreme commander of the Soviet Armed Forces during the Second World War.

The multiplicity of the pieces of information amounts to what Recanati calls descriptive multiplicity. As we just saw in most cases it accounts for the uniqueness of the sense of a proper name, the result being that the name refers to a unique individual. The other, not so frequent, cases are those which the sense is a "one-shot" sense typically characterizing the referent as the first or the only to have done something, or somehow being a distinctive individual (like "the first man in space" referring to Juri Gagarin).⁷ Usually we tend to know quite much about the referents of the familiar proper names. Now I want to use this multiplicity of the senses of proper names to give a Fregean explanation of the differential behavior of the definite descriptions and proper names in the modal contexts. As we recall the direct-causal theorists rely on this phenomenon in their arguments that the definite descriptions (expressing senses) could not be semantically equivalent to proper names. The name "Aristotle" is taken to refer to Aristotle, the famous Greek philosopher (as individuated in the occasion of use), when one talks counterfactually and says that "Aristotle could have been a miserable destitute". But when one says that "the most famous pupil of Plato could have been a miserable destitute" the referent might not be the actual Aristotle, the Greek philosopher, but whoever is the most famous pupil of Plato in the circumstances, in the possible world, the counterfactual talk concerns.

With a proper name is associated an encyclopedia entry which encodes the conceptual (and possibly perceptual) information about the referent of that name and the (neural) representation of that proper name is the pointer enabling the speaker to access that information, it follows naturally that the name tends to activate much of that information in every occasion of the use of the name. That is, a particular use of the name brings with it pretty much the whole informational package associated to it. In

contrast a single description focuses only on a single property of the referent. As individuals have common properties, or may acquire common properties, without being the same individuals, it follows naturally that sometimes the description picks up different individuals in different circumstances. That accounts for the differential modal profiles of proper names and the descriptions. (As the reader realizes this explanation is independent on the actualized descriptions reading defended in the previous chapter; at this point, for a while, the focus is only on the predicative reading of the descriptions.)

Recanati postulates the (in my view narrowly) pragmatic process of synecdoche by which a definite description expressing a "descriptive concept" can stand for a *de re* concept of an individual of which it is a part (Recanati 1993: sec. 15.6). Synecdoche process is a part-whole process in which the informational part engages the whole it is part of. Now it seems to me that this process also helps to explain the differential behavior of proper names and the descriptions in the modal contexts. For it seems, when looked at the direction from proper names to the descriptions, that the part-whole feature involved in Recanati's synecdoche process corresponds to my account that proper names tend to activate larger portions of their senses than the descriptions. It may also be that an activation of a piece of neurocognitive information, expressed by a description, activates automatically the main pointer to that encyclopedia entry, i.e. the proper name of that entry, and the name then activates automatically in a reciprocal manner other portions of the neurocognitive information, other portions of the sense of that name. But because the focus is on the explicitly used description, those other portions are presumably activated to a lesser degree, so that they do not need to become conscious (at least for the time being). So the feature REF, which has the effect that the referent of a directly referential singular term becomes the sole content of the thought, or the singular proposition entertained, does not explain the difference between the modal profiles of proper names and the descriptions; it is the synecdoche process that explains it. From the perspective of my account this is how it should be. It is the different range of the focusing on the neurocognitive information initiated by proper names, and the descriptions respectively, that explains the modal difference. With respect to my encyclopedia entry of Stalin this may happen as follows. If I say "Stalin was the Secretary of the Communist Party of SU" the pointer "Stalin" activates many informationally related pieces, expressible descriptively (though not necessarily all of them). For instance that he was the effective *vozhd* of SU since 1929, the same year when the first 5 years plan was started, and that he worked closely with Molotov. Also other related pieces of information get typically activated but to different extents (depending on the topic). But if I say that "the commissar of Nationalities in 1917 was very busy" the activated pieces of information are presumably fewer and more closely related to the information the description expresses. The name pointer might not even get activated at all, so the certain feel of inevitability that relates to rigidity is not there. The description is more likely to "float free" and therefore assume the feel of non-rigidity *even* if the description is used referentially actualized, hence becoming rigidified. Moreover because the description expresses information that may not be commonly known, it may not be readily inferred that it refers to Stalin. Thus it is easier to see why it tends to float free and tends to have the aura of referential non-rigidity. (But this effect of course depends on what a particular speaker and hearer know.) I hope that this example of the spread of information explicates sufficiently the general mechanism of how sense elements make up the token cognitive significances of proper names as well as the descriptions used.

The situation changes when a description that expresses part of the conceptual information associated with the name is actualized, hence rigidified. In that case the description is equivalent semantically and modally to the proper name; they both pick up the same individual in all counterfactual circumstances. However there still seems to be an "intuitive" discrepancy remaining in this case: it still feels that the rigidified description is not equivalent to the proper name. I suggest that it is the very architecture of proper names as pointers to the conceptual information that explains this residual feeling, along the lines of the Stalin example above. For if the description does not activate all that information at once (though that is possible the sparser the information in the encyclopedia entry is) the other associated information that is not invoked in the rigidification tends to affect the content of the name; it so to speak leaks in. The mechanism is explained in general terms in the next section. In short it is due to the *dynamic* architecture of the encyclopedia entries encoding the senses of proper names that makes it difficult for us to comprehend naturally the result of the rigidification of the descriptions as referential devices when the descriptions are considered at the theoretical level, i.e. not as they are ordinarily used. But of course that does not argue against that very result: proper names and the rigidified descriptions are semantically equivalent.

When a description is actualized it becomes to stand for the whole sense of the proper name the part of which it expresses.⁸ This is equivalent to the role proper names have as the main pointers (as well as conforming to the synecdoche). They so to say collect together temporarily the pieces of the information constituting the sense. In practice this collection is only partial in most cases because usually the senses are quite rich in information. But because of that equivalence it naturally follows that the actualized description behaves just as the proper name does, i.e. it is rigid. The only difference is that any other description as well, expressing a part of the sense, can assume that same role when actualized, but the proper name has that role in all normal referential uses. This of course follows from the architecture of the sense, or the Recanatian encyclopedia entries: proper names frequently provide a wider access to the senses. In every referential use the name functions by being the main pointer and by being activated with the basic schema of reference and this accounts for proper names being *naming* devices. Because proper names (or strictly speaking their neural representations) organize and collect together the neurocognitive information about their referents, they as the main pointers are the usual naming devices. (Though descriptions can become to assume that role also: "The Holy Roman Empire".) The descriptions usually express only parts of the information so their referentiality - the force of REF - is weaker though it is functional all the same. I propose that it is the amount of the information that is being focused on that explains the *relative* strength of REF with respect to proper names and the descriptions. So when Recanati takes REF to be categorical feature I take it to be a graded one and variable in its force in relation to the different kinds of the referential expressions.

The senses and grasping them

As the reader have realized by now we have entered the road to the neuroscientific Fregean account of the senses of proper names. We started with the denial of Evans' categorical distinction of the unbridgeable explanatory levels between persons and information processings in their brains. We then parted company with the neo-Fregean

secretarial analogies which are postulated to account for the cognitive aspects of the senses of proper names. The third step is to look at the dynamic aspects of the senses. That I will do next. The fourth step, and the other main aim of this study, is taken in the next two chapters, in which the Fregean theory is neuronaturalized by the construction of a general neuroscientific theory of the senses and reference of proper names.

So what the senses of proper names are and what it is to grasp such senses, conceived of in the neural-cum-cognitive terms? To answer the first question first, let us concentrate on the cognitive aspects of them (which Frege also thought to be their primary aspects). The senses confer different cognitive significances to the referential expression a speaker uses. Thus one can, presumably, tell who is the referent of the name "Manoel Francisco dos Santos" and also know some facts about a person called "Garrincha", but may still not know, or believe, that they refer to the same individual. Different names can have the same cognitive significances, hence the same senses (apart from the fact that their respective names differ), and one and the same name may have different cognitive significances in different contexts, hence different senses (without being ambiguous, though the speaker may believe so). As I have been arguing information is what the senses are made of, hence different information associated to different expressions, even if they refer to the same individual, explain the different cognitive significances the expressions have for the speaker. Of course the difference have to be such that the speaker does not realize, or infer, on the grounds of that information, that the distinct names refer to one and the same individual; that is why we have cases like "Hesperus"-"Phosphorus". Still, as is obvious, some parts of the senses of two proper names may be the same without that creating such cognitive/epistemic relations between the senses that would lead to the realization that the distinct names have in fact one and the same referent. For example the information expressed by the description "was a superb player in team Brazil back in the 1960s, but later became a drunk" may characterize the senses of "Garrincha" and "Manoel Francisco dos Santos" for a speaker without him realizing that they refer to the same individual. (This account generalizes obviously to other speakers. The senses of distinct proper names, as the information possessed by two (or more) persons, can be so different that the persons do not realize that they are talking about the same individual. But normally persons possess sufficiently similar and cohesive information for them to be able to carry on a conversation with respect to a particular referent. (And of course the cohesiveness increases – or decreases – through such conversational exchanges, i.e. they “measure” the informational similarities of our senses.)

Much of what takes place when a proper name is used to refer goes on automatically and without us being aware of it. Only when a name could not be accessed to be tokened are the descriptions generated from the information storage. This information is the "world knowledge" possessed by the speaker, and it varies from a speaker to another. This informational view of the senses of proper names, and the senses in general, explains naturally the epistemic nature of the senses as cognitive-informational entities.

Let us turn to the second question: what it is to grasp a sense? We are now in the happy position to be able to give an easy answer to that question because the explanation of the nature of the senses above already implies it. This is also good on

the epistemic grounds, for we have achieved a unificatory step. As we remember Frege held the grasping of the senses to be a psychological act. And so it is in my view, albeit a neurocognitive one. One can maintain this by specifying how the senses are neurophysiological entities. But to stay here in the upper steps of our current top-down approach to the matter, the grasping of the senses is in outlines the following type of process. In the first instance grasping a sense is to have those pieces of the neurocognitive information activated which relate to the (purported) referent. This goes on all the time when we engage in a referential use of language. Notice how this account is similar in modern terms to Frege's account. He was primarily interested in explaining thoughts, and they are made up of the component senses of the thoughts. The senses as pieces of the neurocognitive information is what thoughts - or what is also called mental representations - are made of in my account. The only difference is that unlike Frege I do not take the senses to be third realm abstract entities, but physicalistic ones in that they are neurophysiological-cum-informational.

So the grasping a sense of a proper name is to have those pieces of the neurocognitive information activated which relate to that name. (Recall the Stalin example above.) As I explained in the previous section this activation is not the total engagement of the sense; it has parts the activation of which is caused by the information residing in the current context, both the external and the internal cognitive context (like related memories). Of course there are degrees in this type of grasping because the pieces of information one possesses may not be specific enough to enable the speaker to know who or what the referent is. A case in point would be the description "the king of France" which by itself does not activate such pieces of information that would enable one to know immediately which one of the kings of France is the proper referent. But usually the grasp is sufficiently extensive because the context guides the grasp in contributing to the activation of the sufficiently individuating information.

Now we also see in another way why the synonymy demands between proper names and descriptions is empirically unreasonable demand. To establish a synonymy relation would require the linguistic expression of *all* of the sense elements related to a proper name (as a pointer) in question. But that is an insurmountable task in most cases. Only by assuming that there is one or only a few descriptions per proper name is one likely to succumb to the synonymy demand. (Of course the possibility of "massive reduplication" would remain even when all the sense elements have been put in descriptive form; but these I take as irrelevant in the present dialectical situation, for such far out scenarios can always be devised – as any philosopher is well aware.) Moreover if my reconstruction of the notion of individual concept in section 7.5 is correct in outlines, descriptive synonymies are precluded in principle also because the concepts – as senses – involve non-articulatable pieces of information.

It should be mentioned that there is another way to grasp a sense that is a derivative or an extended one. That is to have the sense in one's focus, to explicate it further and to articulate as much as one can those pieces of the neurocognitive information that make up the particular sense one possesses. Notice that this proceeds in most cases at the cognitive and linguistic levels, like when one draws inferences from the pieces of information one already possesses. I am not claiming that there is always such a clear-cut distinction between the two modes of grasping a sense. For there are cases in which the description used turns out to be inadequate because someone else knows more about the referent and points that out to us. In these cases we immediately

become aware of the inadequacy of the sense we possess (or of that part of it which the description invoked expresses). Consequently we modify the sense accordingly during the exchange, and not only, as it were, by contemplating upon the sense in an extended private study. (One might even reject the sense outright in the case it is thoroughly incorrect).

Wettstein's argument

Howard Wettstein has argued that on the sense theory of reference the proposition about the referent a speaker has in mind remains unrevealed in many cases. The proposition is here taken to be the Fregean thought the speaker grasps. Wettstein writes that

"Consider an assertive utterance of 'John was born on March 23, 1976.' If the speaker knows enough about John, he will possess several (non-equivalent) unique characterizations of him. He may think of John as, for example, my best friend, my wife's brother, the best mathematician in Yale, and in countless other ways. For each replacement of the name 'John' in the uttered sentence by such a characterization (or by a conjunction of such characterizations) we obtain a sentence which formulates a different proposition...the speaker...will be often unable to select some one sentence as *the* correct one...If this response is one that we are likely to get, then we cannot suppose that the speaker had a Fregean proposition 'before his consciousness.'" (Wettstein 1981: 150)

I propose that instead of taking Wettstein's point as an argument that purports to show an alleged weakness in the sense theory of reference, it is better to accept it as an *adequate* description of the sense theory, without any of the presumed unwanted consequences. With the preceding characterization of the senses of proper names, as pieces of the neurocognitive information, no explicit mention of the descriptions is required in the normal cases of the referential uses of proper names. Only when the connections between the neural representations of proper names and the information are severed the descriptions become to assume the role of being the devices of the primary referential function (with respect to the particular "lost" proper name). So there need not be any uniquely correct description of the referent in the speaker's mind. Usually there is not except in those cases when the description expresses all that the speaker knows about the referent. Moreover, as I argued, the senses are context-sensitive items. This involves that the components of the senses that get activated depend on the particular emphasis or focus on the referent, for instance what properties and features of the referent are being talked about. It seems to me that this corresponds well to Frege's senses as modes of presentation. This becomes even more obvious when we remember that normally we do possess quite an amount of different kinds of information about individual things, such that we are able to talk about them knowingly (and because of that we also know that others know that the common focus is on the same referent).

There is the difference between proper names and the descriptions in the modal, or intensional, contexts. But as I argued that difference lies in the differential scope of the focus on the range of the information. Proper names tend to activate their senses as

wholes because they are the primary pointers to that information. But it does not follow that the whole sense has to be activated, or that it would be so activated, in order that one would be able to single out the referent of that name. In most cases we do not grasp the whole sense. (Not even in the second sense of grasping a sense, i.e. when it is the object of our cognitive and semantical investigations.) That is the reason why there need not be a full Fregean proposition before the mind of the speaker in Wettstein's sense. But Wettstein's argument fails also in a stronger sense. Such a "full Fregean proposition" (*contra* singular proposition) seems to be *prevented* from being there by the very neurophysiological-cum-cognitive nature of the information residing in the brains of the speakers. This can be explained by the facilitation of those synaptic connections that are involved in the token processes that activate the parts of the information making up the sense in question. In contrast to a proper name a description in its turn activates usually a more restricted number of the elements of the sense. The number of the elements of the whole sense can presumably be explained by the stronger facilitations of the particular synaptic connections (for details see the next chapter). The descriptions so to say focus on a narrower sector of the senses than proper names. So talking about having a full Fregean proposition before one's consciousness is more plausible with proper names, but for the same reasons as with them the content of a description is not restricted to only that component of the sense it linguistically expresses; in many cases the other connections between the elements of the sense, the closely related ones, influence the content of the thought in the particular occasion. Now we also see that the Fregean mode of presentation of the referent is usually not the sense as a whole but it is something in between. From our perspective, that of the modernized Frege, this is how it should be (as I remember telling a couple of times before): the senses are the informational elements of the thoughts and the linguistic expressions are the resulting condensations of them.

Because of the differential focusings on the senses, either via the name or via a description (expressing directly "its" part of the sense of the name) we can now further constrain Recanati's notion of encyclopedia entry. For we see that it is only natural to modify the "democratic" architecture of that notion. (To remind the reader, it is neither democratic in the sense that all speakers would possess the same entries.) That is, in Recanati's original account the entries of an individual encyclopedia are all on a par, and it is not clear what could restrict all of them from being activated in every occasion of access of the entry in question. But if we suppose that via the name as a pointer one gains access to 1) some pieces of information, expressible by descriptions that are about the properties of the referent that characterize it in many cases and/or to the most degree for the speaker; and 2) those pieces have their own subpieces of information, expressible by the descriptions relating only to the former descriptions; and, which is likely, iii) these divide into further subpieces, then we can have a more adequate architecture of the encyclopedia entries as the senses of proper names than we can have with only the "phenomenological" means the neo-Fregean approach allows.

This architecture can be represented by a tree structure some branches of which are connected to one another between the trees. In other words some pieces of information characterize many referents. This property of the senses, as Recanati's encyclopedia entries, accounts also for the occasional fluent engagement of the information about other referents, be they persons, events or objects that are related to the primary referent in the occasion. That explains the notion of the "surrounding story" I am using

in some places in this work. That spells out the phenomenon that the referents of proper names are readily located within larger informational contexts, the token representations of parts of these contexts becoming parts of the senses of proper names. Hence it accounts also for the contextuality of the Fregean senses that we detected in section 4.3.5. To this property of the senses adds what I noted above, and what also the Ateb-Afla example implies, that the entries can contain perceptual information as well and not just conceptual information.

We can also explain the mechanism underlying the use of the actualized descriptions. When a speaker says “the ninth chess World Champion was a cunning positional player” he refers to Tigran Petrosian and intends tacitly that this use subsumes also the ordinary counterfactual contexts (“the ninth chess World Champion could have beaten Spasski if he had put more effort to his play in 1969”). That is because the description activates other pieces of information related to it – it activates those neural sites wherein information about and related to Petrosian resides. Primarily this includes knowledge about the history of the top class chess, Petrosian and games he played. And these in their turn tend to activate the more indirect pieces of information – presumably more weakly activated neural information (senses elements that become to assume a “stand-by” state) – related to the former informational sense elements about Petrosian in the speaker’s overall sense system.

It seems that the surrounding sense elements explain also at least part of the “intuitiveness” of the direct-causal theorists’ arguments like that of Feynman case. Because we typically have plenty of the elements related to proper names we use, and we know how to use the names to refer to a particular individual, it becomes automatic to judge that despite the lack of individuating information in some cases the names in question still refer as usual. Hence it seems that “Feynman” refers to Richard Feynman the physicist in the ignorants’ mouths. That is, we possess amply individuating information, direct and indirect, about the referents in the majority of cases and that could not but influence the interpretation of the Feynman cases where there is no individuating information possessed *by a particular speaker*. Because we know that i) we usually know who we are talking about *and* that ii) if one does not know, there are others who know and to whom one could depend on by deference, it automatically seems that the proper name does refer irrespective of any individuating information. But, to repeat it, the semantic reference as deference is here confused with genuine reference: the cognitive information-based reference by a particular speaker (or speakers when all they know who the referent is). That the “intuition”, or feeling, of secure reference in the Feynman case is not to be trusted as such is evident from such cases where one believes that there are experts who know the referent, but when in fact there are not for the simple reason that the purported referent does not even exist (unbeknownst to all the speakers). So the “intuitions” reflect only the general conditions of the referential use of language, they do not reveal anything substantial about genuine reference. It is easy to see that this same explanation applies to the Gödel cases as well: despite the sufficiently individuating information involved, the intuition that “Gödel” refers to Kurt Gödel, not to Schmidt, taunts us because we know “too much” about Kurt Gödel. And when it taunts someone who does not know, that speaker’s verdict in favour of Gödel is explained by the trust on deference.

Concluding remarks

In concluding this chapter it should also be pointed out that the above explanation of the senses and the grasping of them help us to explain why the notion of grasping has remained rather unclear so far. Because the senses are processual neurocognitive entities there is nothing surprising that the earlier accounts of grasping have remained so timid. Also the tendency – characteristic of philosophers - to search for distinct and linguistically wholly articulatable notions has led us to demand of the senses more than they can deliver because of their very ontological nature. I think that the characterization of the senses of proper names as neurocognitive informational entities shows also that it is a high time to retreat from the overly linguistic conceptions of how language functions referentially.

It could not be denied with any plausibility that the cognitive level studies, at which the earlier neo-Fregean studies have ventured, are important by themselves and as the forerunners, albeit inadvertent in most cases, of the correcting themes to merely folk psychological and folk semantic conceptions - though this aspect of the neo-Fregean studies is rarely explicit because they still tend to assume the basic explanatory soundness of those frameworks. However many cognitive scientists have also stopped at the cognitive level and have deemed the further attempts at searching for the neural level accounts as irrelevant. The common opinion seems to be that the neural facts only "implement" the cognitive ones, or only "realize" them (or that the latter "supervene" on the former). It also seems to be assumed that the search would not end up with any adequate theory because of the sheer complexity of the brain's structures and processes. These opinions I take to be true as to the complexity but false as to their suggested negative implications: we might or we might not end up with explanatorily adequate neurocognitive theories of reference, but we do not know that beforehand. Here as elsewhere when we are taking the first steps toward explanations in a new domain opening up the burdens do certainly look considerable, but that is no reason to refrain from attempting. I will not engage in arguing about the ontological issues of conceding mere implementation or some such role to the neural level. As the next chapter makes it plain I take it as a fact that the neural does not merely implement the cognitive but it *is* the cognitive, even contentwise. This is why I take the senses to *be* neural entities informationally, i.e. the information is literally there *in* the neural processes, and the cognitive level phenomena and entities emerge from those processes informationally.⁹

Appendix: Wettstein's anti-Fregean revolutionary semantics

In this appendix I will concentrate on a rather curious twist the direct-causal consensus has fertilized. The background is that the approach has repeatedly failed to explain the famous "Frege's puzzles" of the cognitive significance of proper names and other cases. These puzzles center on what has also become to be called "Frege's data" (Wettstein 1986: 186). Now I take it that Frege's data is the *basic data* any theory of semantics and reference must explain. Therefore it is remarkable that the constant failures of the direct-causal theory of reference in accounting for that data have as constantly not been let to bother the proponents of that theory. The curious twist I alluded to is that, in addition to this indifferent attitude itself, one writer agitates for a further dismissive perspective on the data; indeed for a revolution in philosophical semantics. That writer is Howard Wettstein and the propaganda for the revolution has

its beginnings in one of his paper from the mid-1980s (Wettstein 1986) and is collected together in his book *Has Semantics Rested on a Mistake?* (Wettstein 1991).

Wettstein argues that we should give up being bothered by Frege's data in referential semantics. In effect Wettstein states that we could ignore it and be no worse off; cognitive significance is a real phenomenon, but it has no relevance to referential semantics. This is the revolutionary impact of his perspective. In the following I take a critical look at some of the most important of Wettstein's arguments and ideas. In particular I will argue that Wettstein's view is not sustainable in that it is empirically inadequate even on its own terms.

First of all it should be noted that unexpectedly Wettstein's understanding of the Fregean theory is that it is a descriptive theory. According to Wettstein the senses of proper names, in general the referential expressions, are fully expressible by descriptive means. Wettstein's arguments against the Fregean theory are premised on that conception, and that results in them possessing rather limited value, for the Fregeans need not, and do not, claim that the senses are wholly descriptive. Moreover Wettstein works also under another widespread misconception that the Fregean senses must be purely qualitative. According to him "Frege advanced a view...that a name user need attach to the name a purely qualitative concept that he takes to single out a referent..." (Wettstein 1991: 156) As I argued this is also an incorrect assumption, but for the sake of the argument I ignore it and deal with Wettstein's arguments as they are.

Wettstein contrasts two conceptions of semantics, Fregean and anti-Fregean. As mentioned this involves the wholesale denial of "Frege's data" as relevant for semantics (Wettstein 1991: 27-8). But right here in laying out the central issue lies a serious problem. The denial of Frege's data means that semantics proper should not have anything to do with the phenomenon of cognitive significance, or cognitive content, of the referential expressions. It should be only the notion of reference as such that semantics is concerned with, and the actual practices of the speakers creating and using referential expressions. All the cognitive stuff, the modes of presentations of the objects, the ways of thinking about the referents, beliefs about them, and such are mere paraphernalia well ignored in the semantics proper. This is revolutionary indeed. The real things on which Wettsteinian semantics advises us to concentrate are the conventional and institutional forms of communication with the referential expressions. In this he likens his view with Wittgenstein's studies of language games and their roles in the linguistic communities.

But we must not be taken in by the charm of the revolutionary aspects Wettstein's propaganda advertises. That is because – as we have been seeing - the fact remains that cognitive significance and the modes of presentation, i.e. pieces of information about the referents, *are* so intimately tied to the referential expressions that it seems singularly implausible to even to try to sever them from reference without committing gross *ad hoc* moves. For instance the identity puzzle *is* about the modes of presentation of the referents. The importance of that puzzle is that it shows, *because it is generalizable*, that every referential expression has a mode (or modes) of presentation of its referent. Wettstein is very hard put to deny the relevance of the identity puzzle because it manifests itself in the everyday referential practices, i.e. the very data of Wettsteinian "semantic anthropologist". In other words, we do not use

proper names alone but express information and beliefs related to them; we always talk about individuals by saying that they are such and such and did so and so.

It goes also the other way around. When we are reduced from exchanging information about a referent, for instance due to the information being all incorrect or due to the sheer lack of information about a historical figure whose name we only know, the name will before long drop out of our institutional practices if we will not acquire adequate information about the bearer. (That was argued in chapter 4, so I propose to take it for a fact now.) Consequently that figure will disappear from the scene of the history with the name. The only historical figures we are currently talking about are those about whom informational records have survived, mostly in the form of literary sources, tales and the beliefs of both ordinary and expert speakers like the historians. All other persons have faded into oblivion for some reason or another. For example one will quickly realize this when backtracking one's family lineage -though if one is a descendant of a prominent figure that will take longer depending on how far back in the generation count that figure is to be found. This general phenomenon should be of interest to the *anthropological* semanticist. But that very phenomenon reveals that it is explained only by focusing on information about the referents, of the modes of presentation, and these bring with them the phenomenon of cognitive significance back again.

Of course one can adopt synchronic view instead of diachronic. One can argue that the anthropological semanticist is more interested in the contemporary referential practices, suitably restricted. But how would that help to evade the argument from information? Maybe Wettstein is likely to argue that it does not take much effort for anyone to get locked to a proper referential practice with a proper name; for instance that could happen if one merely overhears something being said about someone with the explicit use of the name of that target. (Sometimes the direct-causal theorists indeed have claimed this.) But we must look at this case more closely. First, in conformity to what I pointed out above, with these overheard uses information is still transmitted in the overwhelming majority of cases. Second, when one's epistemic resources, one's beliefs, about the referent of that name are very scarce and not capable of enabling the speaker identifyingly to refer to the bearer of that name, to discuss about him, what to do? Here Wettstein can call the direct-causal theory for help and state that one is still locked in the appropriate referential practice with respect to that name because one has become part of the overall referential network which includes experts, i.e. those who can identifyingly refer to the bearer. But I am sure it will not escape anyone's notice that here the ignorant speakers are again made to refer by deference, and the success of that mechanism depends on the information the experts possess about the bearer of the name in question. If the ignorants were to be left alone, their referential helplessness with respect to the bearer of the proper name in question would soon manifest itself (excepting trivial hypotheses like he is a human, and such).

It seems that Wettstein has overlooked the direction of explanatory dependency. The referential conventions and institutions on which the anthropological semanticists dwell are sustained only *because* people talk about things and other persons by exchanging information. They do that always synchronically (how otherwise could they do it?); but from the synchronic events diachronic dimension grows. That is, the referential conventions and institutions supervene on the synchronic practices. To keep

one's semantic eye only on these general uniformities – theoretically cleansed from the information that really is what makes them effective - is going to the diametrically wrong direction explanationwise. At best Wettsteinian semanticist only describes (parts of) the typical patterns of the use of proper names.

The facts of the referential practices being what they are, i.e. based on information transmission, it is incorrect to claim that the Fregean theory is incompatible with the actual linguistic practice (Wettstein 1986: 201). This observation can be reinforced when we take a look at one of Wettstein's own main examples (Wettstein 1991: 150). Suppose that someone has beliefs that apply to Socrates (Plato's teacher, philosopher who drank hemlock) but uses the name "Aristotle" all the time. In relation to the conventional and institutional means of using referring expressions it is somewhat ironical that Wettstein stops short in his presentation of this example and only claims that there is no reason to prefer the Fregean view according to which the speaker is talking about Socrates, instead of the "New Theory of Reference"-inspired view that he is talking about Aristotle. The irony here is that, if the anthropological semantics really studies the social patterns of referential communication, Wettstein should have looked at what will happen in the possible subsequent situations when the speaker keeps on using "Aristotle" while having beliefs applying only to Socrates. If the speaker has managed to get along with his uses of "Aristotle", i.e. he has not said anything that would have revealed the discrepancy between his uses of "Aristotle" and his beliefs related to that name, he has not become under any suspicion of referential malpractice. But presumably that is only because his audiences have no discrepancies between their beliefs about Socrates and their uses of "Aristotle" (and "Socrates"), due to their accurate knowledge of the history of philosophy. But when the speaker will say things like "he was the teacher of Plato", he is obviously taken to be referring to Socrates by the expert audience. But then when he uses "Aristotle" and say things that are true of Socrates only, like "Aristotle drank hemlock", he will be corrected and his use of "Aristotle" made to conform to the prevailing referential practice. We see that what is said weighs much more than about whom it is purportedly said. But this is just what the approach relying on the primacy of the information transmission predicts.¹⁰

The next problem, related to what was just argued, is that Wettstein generalizes even further his non-informational anti-Fregean account (though, of course, he lets us know this is a very positive thing; and how else it could be revolutionary). Wettstein claims that it is the loose linguistic practice which in most cases does not have, or even allow for, any "cognitive fix" such that would keep the users of the referential expressions tied to the referents of those expressions. He notes that

"the crucial question is whether reference is to be tied to the community's *beliefs* or rather to its *practices*. Couldn't it be, for example, that our current beliefs about a historical individual have become all fouled up but that the continuity of usage secures reference nevertheless? Our beliefs about *him* are, in such case, mistaken, but there remains an individual about whom we are talking and about whom we are mistaken." (Wettstein 1991: 224-5, n.19)

In addition to the negative answer, already given to Wettstein's innocently looking question about *him*, this quotation gives immediately rise to some other crucial questions.¹¹ Why Wettstein thinks that beliefs and practices can be distinguished for

referential purposes, and for explanatory purposes, in the strict way that he assumes? Or that they are in fact so distinguished? It does rather seem that the practices, the conventions of using proper names and other referential expressions, have become to be the way they are because our beliefs of both the referents and the world in large have been what they were and still are? Before these questions are given such answers that would support the beliefs-practices distinction, it is advisable that we do not jump to Wettstein's revolutionary bandwagon even on this count.

Although the preceding discussion and the argument are condensed, I tend to take them conclusively showing that in propagating the anthropological semantics which studies the conventions and rules, and whatever, of the referential practices of language use, Wettstein's account subtracts the very things that keep the conventions operative. Wettstein claims that "there is no reason to suppose that, in general, if we successfully uncover the institutionalized conventions governing the references of our terms, we will have captured the ways in which speakers think about their referents." (Wettstein 1986: 201) But this claims too much for it is not demanded by the Fregean approach that the indefinitely many particular ways of thinking the referents should be captured. It suffices to explain, in general, how the ways of thinking about the referents give rise to the institutionalized conventions and practices governing the references of our terms. And that is what the senses as cognitive fixes do.¹²

Another weakness in Wettstein's argument is his treatment of "that" (Wettstein 1986: 202-3). He claims that any of the plausible explications of the rules by which the expression "that" refers (Kripkean causal relation, the object the speaker is having in mind, the contextual cues - which is Wettstein's own account) do not have anything to do with how the speaker thinks about the referent, with what is his cognitive fix on the referent. But Wettstein could only make it seem that his argument supports the anthropological semantics because in this connection he ignores proper names. It is easier to try to make the case stick with demonstrative expressions like "that", for even to the Fregeans the cognitive content of such expressions is rather meager. But the demonstrative "that" *does* have cognitive content, for how else could it be different from the demonstrative "this"? That is, it is mainly the spatial distance from the speaker to the referent what decides whether the speaker uses "that" or "this". And the spatial cognitive fix *is* a plausible Fregean sense element. (And one is hard put to deny that it is such after Evans' studies (Evans 1982).) (Though there is also the meaning relating to the order of comparison: "this is larger than that".)

The foregoing argumentation concerns the cognitive significance of simple, unembedded sentences. Wettstein shortly discusses embedded sentences and the propositional attitude contexts in the end of his paper (Wettstein 1986). The main problem of the anti-Fregeans is well known as was already noted in the beginning of this appendix. For instance if one believes that Cicero was an orator, and assents to the sentence expressing that belief, one still might not assent (and could even dissent) to the sentence "Tully was an orator". But because Cicero was Tully this should not happen according to the (Millian and neo-Russellian) anti-Fregean view: the semantic values of the two sentences are the same singular propositions. The Fregeans have no problems here: the senses of "Cicero" and "Tully" differ from each other, so the truth-values can differ in the propositional attitude contexts.

Wettstein emphasizes that in this case the data is *not* to be ignored or denied. That is, it is not any (mis)solution to maintain that the truth-values of the Cicero-Tully sentences in the propositional attitude contexts do not really differ. However he claims that the consequences of that data may "not be what they have seemed" (Wettstein 1986: 205). But it is difficult to see what relevance Wettstein's proposed way of looking at the attitude/embedding contexts could have with respect to the unitary explanatory force of the Fregean approach. The reason is that the direct-causal theory of reference, and the singular propositions wedded to it as well as Wettstein's own attempt to replace the propositions with states of affairs, are all unable to handle the attitude phenomena to begin with because these postulations are extensional. *Cognitive* significance just is a fact of the referential language use, and a fact that by its very nature is incompatible with any coarse-grained proposals that work only with the extensions and the referents of the expressions involved. Therefore I maintain that it will not do any good to deliberately try to ignore it the way Wettstein attempts.

That there are some problems with the Fregean way is of course not to be denied. But they seem to be not sufficiently serious to threaten the basic explanatory structure of that view. Wettstein argues that we often do not even try to retain the Fregean thought the believer grasps when we report what he believes. Wettstein gives everyday examples in which proper names are replaced by other proper names and descriptions. This is fine as far as it goes. But Wettstein so to say over demands from the Fregean view what it is not committed to. There is no need for the report of the belief to re-express exactly the Fregean thought. I would even propose that the assumption from exactness is why the attitude reports have been so difficult to handle so far (see section 7.1.3). So even if I concur with Wettstein on this point, I need not concede that this brings any threatening consequences to the Fregean approach.

Wettstein also argues that the attitude reports involve social and contextual factors to such an extent that the semantic analysis of the attitude reports is even messier than has been thought before. I think that this is basically correct but, once again, nothing negative with respect to the Fregean approach follows. The consequences are quite the contrary, in fact. For the social and contextual factors that has to be taken into account when reporting beliefs (and other propositional attitudes) only add to the variance resulting from the different cognitive fixes people have of the referents of the expressions they use. And as I argued Frege's own account allows for such factors. Neither did Frege commit himself to the reporting of the strictly same thought. This follows from his view that different persons may well possess different senses, hence thoughts, related to the same referential expression. This is what Frege said in the (also in this work) often mentioned "Aristotle note". It naturally follows that the beliefs, the Fregean thoughts, do differ to some extent between the believer and the reporter.

So all in all, the Wett(gen)steinian revolution is not likely to take place any day soon - as it has not since Wettstein's book was published. Note that I am not implying that the study of the social practices, institutions and conventions of the referential uses of language are not useful explanatorily, let alone descriptively. But these studies focus on the "superstructure" of reference, on the wider and more general patterns primarily emerging from the "microlevel" facts of language use that only the sense theory has the full potential to explain. Reference with proper names succeeds not because the names are external to the speaker's cognitive apparatus and are social instruments, but because as such they are the results of the internal informational factors we possess

individually. The social patterns and practices result from the co-operation and informational “negotiations” in cases of referential conflicts, which help to link our respective pieces of information.

Of course Wettstein’s vision might not be a dream impossible to make true. But the only possibility I see for its success would require very drastic changes such that Wettstein has not even hinted at; the mere “radical” dismissal of Frege’s data and the related cognitive facts will not do. Let me explain. Wettstein writes, in a fairly partisan tone, that “Freed from Frege’s perspective, we will, I contend, no longer find it natural to think about semantics in this way. And so it will no longer seem a decisive objection to the anti-Fregean semantical work that no immediate solution to the cognitive significance puzzles falls out of that work.” (Wettstein 1991: 8) Wettstein likens this to Wittgenstein’s idea of dissolving philosophical problems and deep-entrenched views. (Though that is not especially Wittgensteinian activity: much of philosophy, but also science, is such.) What is important is that it should be realized that this strategy of dissolving commits one to the changes of subject, to the rejection of the previous conceptions of semantics. Now this *is* what revolutions are mostly about, but Wettstein is then committed to non-argumentative revolutionary change because his arguments fail (along with those the direct-causal theorists have offered). The very cognitive nature of our thinking and language use is so essential to semantics as we have it and *live* it in everyday use of language that no argument could change that. We should really have to be *freed* from the Fregean perspective, and that would require very different methods than the cozy arguments. But what are those methods, then? In effect I am claiming that the cognitive fix, via the ubiquitous informational links and our biological constitution, prevents us from having any other semantics than the Fregean. (The developments in the next chapter should be taken as providing a general argument for that claim.) This makes sense as a sort of scientific transcendental argument, because the biological facts are more basic than the anthropological ones. For one thing, obviously, our anthropological practices and uniformities are constrained by our biological nature. Consequently to be freed from the Fregean semantics would not just involve being freed from a *perspective*, but being “freed” from our very *biological nature*. In short, only by being changed into a very different species would we achieve any non-Fregean semantic revolution Wettstein dreams of. But I doubt that that will ever be possible in our kind of world because without the informational traffic between an organism and its surroundings the former is unlikely to survive. So it seems that Wettstein’s envisioned program is not a form of semantic revolution but ultimately a form of elimination of all semantics.¹³

In Wettstein’s recent book from the year 2004, *The Magic Prism*, very little has changed argumentatively. Wettstein still mentions – as has become the usual practice of the anti-Fregeans of all kinds – approvingly the Kripkean arguments against the Fregean approach. But as we have seen these arguments fail, so there is no need to go all over that again. Wettstein’s only new argument for his variety of anti-Fregeanism has gone even more on the tangent: he claims that the Fregean modes of presentation of referents do not explain the informativeness of the identity sentences. Consequently Wettstein is ready to abandon the modes of presentation in semantic studies altogether (“...modes of presentation are a bad idea...” (Wettstein 2004: 135)). In short Wettstein’s argument is the following (Wettstein 2004: 134-8). He invokes Putnam’s example of the concepts of *elm* and *beech* which Putnam says he cannot tell apart: for him they are both expressed by “the large deciduous trees that grow in the east”.

(“Elm” and “beech” are supposed to be natural kind terms of course, but let us that pass here, especially since Wettstein remarks that Kripke’s Feynman example is similar to this case.) So there is only one mode of presentation associated with both concepts. But if someone tells Putnam (as is obviously the case) that elms and beeches are different trees, Putnam would acquire new information. Wettstein concludes that informativeness/cognitive significance does not require different modes of presentation, hence they do not determine reference. That is, Putnam knows that elms and beeches are different but still the modes of presentation have remained the same for him.

The Fregean objection to the effect that the talk of the modes of presentation *is* talk about the information a speaker has at his disposal, and could supply if asked, Wettstein brushes aside by claiming that “it is neither here nor there that, in some cases, she [the speaker] had the information available to her in making the original reference.” But I fail to see what could be the point of this reply. If the speaker could supply the *individuating* information by whatever means that will suffice for her referential purposes. In other words Wettstein makes the same mistake as Kripke in that he thinks that any mode of presentation will do. No, only individuating modes will do, but “large deciduous tree that grows in the east” is not such. Wettstein tries to counter that *often* the speakers do not have sufficiently individuating information at their disposals. This is, again, one of the typical claims the anti-Fregeans keep on repeating time and again. But to my knowledge it has never yet been supported by any empirical studies, only isolated anecdotal examples like Einstein being the inventor of the atomic bomb. And as I have mentioned the contrary seems to be the case: we often do have individuating information about the referents (and extensions). If we did not what could be the point of using proper names? How could that practice have survived without any individuating information related to the names? Of course there are also cases when the users do not have individuating information, barely any information, about the referents of proper names they use. But the anti-Fregeans somehow take it that this automatically shows that proper names still genuinely refer. This I take to be plain *non sequitur*. To repeat my argument: the anti-Fregeans confuse the semantic reference by deference with the speaker reference, i.e. they take the metasupposition (that when a proper name is used there is a known referent) for genuine reference, and so confuse a precondition for substantial referential use of a proper name.

Related to what was just argued and to round off the critique of Wettstein it seems to me that it is the referential preconditions, deferential metasuppositions most prominently, that Wettstein in effect means when he speaks of “linguistic anthropology”. But how curious that is, since 1991 Wettstein has not produced any case study of this anthropology; only propaganda in very general terms and an almost naively sincerely looking belief that the anti-Fregean arguments are correct. In light of this I think that I can conclude that Wettstein’s revolution is but a misapprehension.

Notes

1. As Evans emphasizes the verificationism inherent in Dummett’s notion of sense is of an ideal type. Senses do involve a verificationist element because they are the modes of presentation of the referents by which we “aim” and “reach for” their targets. But to find out whether the target is such and such is another matter and requires stronger means of verification. Dummett also takes part of the explication of the senses to be their roles in

finding out the truth-values of sentences. But this seems to be much too liberal, for it does not preclude whatever specific means one could use, even accidental ones. Evans noticed this also (Evans 1982: 96-100).

2. This adoption of the categorical distinction between the subject and the brain processes is apparently a case of the "Oxford syndrome", the roots of which derive from the misconception (used in argumentation back in the 1950s) that the talk about mental states differs ontologically from the talk about the brain states (because the former are not spatial in their "logical syntax" whereas the latter are). John McDowell exhibits this syndrome when he states that postulations of implicit knowledge/information structures of mind, in order to help to explain language use and understanding, really do not shed any scientific light on those phenomena; rather that brings "philosophical darkness" (McDowell 1977). I am astounded. How do scientific attempts at explaining bring philosophical darkness – unless the very philosophical conception of explaining is incommensurable with scientific one? If so, that I simply take to constitute sufficient ground for a *tollens* argument, and such that amounts to a *reductio* of the Oxonian philosophical conception of explanation. I regard it hardly worth of anything to elaborate more on the issue. This kind of Oxonian apriorism just will not do, if only for the fact that – as McDowell himself admits, language understanding is what one has to explain with the Fregean senses – the Fregean senses are *cognitive* postulations (they have cognitive significance, *Erkenntniswert*) and cognitive postulations (including "informational structures") *can* be studied scientifically. Mere philosophical explanation is not sufficient; it is too bound to remain too high-level enterprise by itself. To the extent that this involves "psychologism" that Frege detested (as does McDowell), the issue is much too coarsely characterized by McDowell. The point is the interplay of subjective cognitive information structures and mechanisms and how the exercise of them brings about the objective (intrasubjective) patterns of referential language use and understanding. I take it that this bringing about (akin to psychosocial emergence of the patterns of language use) of the intersubjective factors is an obvious fact. "Psychologism" is a red herring here: the validity and soundness, so to say, of the patterns can be explained by the subjective-cum-biological factors like the need to cohere with others. Here philosophy without neuro and cognitive sciences is a sort of theoretical Parable of the Blind.
3. As to Evans' understanding of Frege's view I think that Evans puts unnecessary weight on what he calls Russell's Principle, according to which no thought has been entertained if the referent of a singular term, in the sentence or an utterance which purportedly expresses the thought (or should it be "thought", or schthought if Russell's Principle is heeded?), does not exist (Evans 1982: 43-4). In some places Evans' arguments reduce to mere claims to the effect that there has to be an existing object for there to be genuine communication at all (see for example (Evans 1982: 336-7)). I can not help thinking that the whole strategy of extending Russell's Principle to the class of singular terms, expressive of particular thoughts (Evans 1982: 109, n.32), is just an instance of the mistake of the idealization of the causal connections from the demonstrative references, which the direct-causal theorists have made. Indeed, demonstrative identification of the referent (of the thought) plays a prominent role in Evans' argumentation. (And it was no coincidence that Russell countenanced demonstrative and indexical expressions as referential, as the only logically proper names.)
4. Evans claims that there follows an apparent inconsistency with respect to Frege's theory of sense and reference when it is conceived along the lines that a sense is given when the

referent is given. For if there is no referent, the singular term is empty, and there is no sense. But Evans corrects the apparency. For Frege such cases belong to the realm of fiction (as he did say). In that realm the senses and thoughts are expressed, but what is thereby said is fictional just because proper names in question do not have real referents. Frege also called such thoughts mock thoughts (*Scheingedanke*). Evans interprets this in the way that the thoughts expressed are only pretended to be thoughts. But that seems incorrect: mock thoughts or not, they *are* thoughts. It seems to me rather that "fiction" was an umbrella term for Frege. It includes all thoughts and senses that have no real *Bedeutungen*. Evans takes fiction too literally, so to say. This mistake haunts also his whole explanation of the negative existentials by pretense attitude (Evans 1982: 353-68). The explanation is not just *ad hoc* but goes contrary to the facts: we do use empty names when we do not know them to be such, and we do this without any pretense (tacit or not). This is fundamentally damaging to Evans' Russellianism because he states that the pretense explanation of the negative existentials, and other such types of (seemingly) non-Russellian linguistic phenomena, is the only one compatible with Russellianism (Evans 1982: 340). (This same point, though with a different emphasis, is made by David Bell when he criticizes Evans' interpretation of Frege's mock thoughts (Bell 1990).)

5. I do not accept Luntley's theory of sense "as the way our behavior is rationally structured" (Luntley 1999: 281). Here I do not have the space to develop the arguments in detail, so I only note the main points of disagreement. Not even our everyday behavior is uniformly rational, and to account for its irrationalities and quirks the senses are also needed. Focus on the behavior ignores the mass of the internal processings by which a thinker-cum-speaker makes sense of his surroundings, forms new hypotheses and infers in ways that changes some of the other senses he possesses. Consequently Luntley's account is too narrow. But it is also too broad because it does not seem to be properly Fregean at all in its rather general emphasis on the external information, in comparison to how the thinker-cum-speaker utilizes it, i.e. in relation to the cognitive and epistemic functions the senses make possible.
6. In the case of the descriptions and proper names the exemplar/typical representative is trivial: the sole and unique referent of the expression. But this of course does not militate against the point of perceptual information being part of the dossiers (because not all of that information can be descriptively expressed). There are also other open questions, though ones that do not threaten the basic architecture of Recanati's account, for example are there distinct representations, one for each individual encyclopedia entry, of the shared properties of the referents the thinker knows?
7. There is a connection to what I will call ontological, or Leibnizean, individual concepts in the next chapter. If one could be able to guarantee that an individual sense determines uniquely its referent, one could add on and on the facts about the referent to the "dossier", the secretarial analogy of individual concepts, thereby approaching the Leibnizean complete individual concept. This does *not* help against the metaphysical possibility of massive reduplication (or such extreme Twin Earth scenarios). But one does not have to worry about these anyway, for as I argued they are out of the bounds with respect to what our cognitive and epistemic capacities allow to be reasonable considerations in semantics and theory of reference, taken synchronically. Moreover invoking reduplications can be used against the direct-causal theory as well. One can devise a scenario with mixed causal links such that on the direct-causal premises there is no fact of the matter what the

referential term picks up. (To call Field-Devitt partial reference for help is to miss the mark, for that notion is as allowable for the Fregean and the description theorists.)

8. This is synecdoche process. Recanati states that it is a property of *de re* concepts of individuals (proper names refer to) that a description can assume the role of synecdoche. But this seems to me too restrictive. The same process can be effective in the purported uses of proper names (via the descriptions) in the case of empty names as well. That is, from the perspective of synecdoche there is no difference to the normal uses of proper names that do not refer to any existing individuals unbeknownst to the users – which is why they are used normally.
9. In the appendix to chapter 4 I criticized Devitt's account of non-Fregean senses as causal chains residing partly, or mostly, inside the brain of the speaker. But am I not committing that same mistake when I claim that the senses are neural entities ontologically? No because my account is different. I identify the senses with the neural network processes containing *information*. It is from these processes within which the neurocognitive information specific to the referents *emerges* as neural. The basic ontological and explanatory difference resides in the property that those processes can access one another because of their informativity (i.e. because of their content, their content features) whereas the causal networks Devitt takes to be senses are mostly out of the range of the speaker's access in that way. And those parts, the chains and the networks in his brain, to which he has access are just the informational processes from which the senses emerge – or which the senses *are* – in the interplay of the received information and the already possessed information. So in effect Devitt's account does *not* differ from the Fregean one. I am not sure that it characterizes the difference totally, but I surmise that it is the degree of informational complexity *through the processing architecture* which makes for the difference between the non-cognitive and the cognitive relevance, i.e. the cognitive emerges from the causal-cum-informativity through the brain's dynamic architecture. No such comparable feature is related to the causal links and networks outside the speaker's brain, so the information they contain remain insufficient to attain cognitive status. (It is not utilized *in* those outside networks.) It should be kept in mind that Devitt's notion of access is the culprit here. For him it is sufficient to have access to a part of the network to belong, *qua* speaker, to the causal network. For me access requires that one is able to utilize part(s) of the network. In other words access requires acquiring information and being able to process it in one's cognitive apparatus. To put it picturesquely, only when the causal folds unto itself do the cognitive features and properties emerge. (See also note 1 in chapter 7.)
10. There is not a fixed fact as to whom the speaker refers when he uses "Aristotle"; it depends on in what "informational surrounding" he uses it, and in what way he uses it even in that surrounding. Wettstein acknowledges this, or something similar to it (Wettstein 1991: 150-1). But he fails to realize that the referentially deciding role in those situations falls upon the individuating information, not upon proper names. To claim that the speaker is always referring to some definite individual, then, is only a demand that is cherished by Wettstein along with the direct-causal theorists of reference, but one that the facts of the very communicative practices do not respect in most cases. (Wettstein also realizes that there may be no one thing thought about and referred to (Wettstein 1991: 218, n. 41). But even here the urge to have a referent sweeps aside the reasonable alternative that the speaker possibly does not refer to anything.)

11. The whole answer being that once we discover that the beliefs are fouled up we are very likely to give up on using the name, or we use it only very tentatively to refer to *him* because we are not at all sure that it does refer to someone. (By the way, how can Wettstein say that it is *him* we are referring to if we do not know anything about the target?)
12. It is partly correct that the conventions will not help with the questions of the speakers' cognitive perspectives (Wettstein 1986: 201), but that is because Wettstein looks at the issue from the wrong end, and also does not realize the general way to explain the conventions as emerging from and supported by the particular referential uses of language.
13. This observation can be generalized further. If our biological nature deems us to the Fregean semantics in the way I argued, it becomes also arguable that the Fregean semantics is operative with *all* biological species. By this I mean that all biological organisms have to survive and adapt in their respective (but mutual) environments, their overlapping ecological niches. For all we know the overwhelming majority of these species lack language-like communication (if physical and chemical signals are excluded from what can constitute language). But they still survive and adapt, which I like to construe as supporting the thesis that they, or at least those species with some kinds of nervous systems or ganglia, have very rudimentary *senses*, maybe even rudimentary *concepts*, of the objects and events in their surroundings, therefore Fregean graspings of them. (Remember that Frege himself allowed that for grasping senses language is not necessary.)

7. AN ATTEMPT AT NEUROCOGNITIVE THEORY OF REFERENCE

On methodology

I would like to begin this second part of my attempt at the neuroscientific framework for the reference of proper names, and reference in general by obvious intention, by restating the point about the methodology I made in the introduction. The neuroscientific, or rather neurotheoretical, perspective might seem to some diehards to be out of place in a work that is supposed to belong to philosophy. That I could not accept. My motivation for the neuroscientific approach is very philosophical indeed. As I see it no adequate explanatory account of reference could be achieved without taking the relevant sciences into account. It is a metaphilosophical motivation in that reference research has been too much in the grip of old-fashioned "philosophical" methodology (like counterexamples to necessary-and-sufficient-conditions analyses, demands of strong semantical equivalences). This manifests a form of naturalism on my part: no strict wedges of any kind between philosophy and the special sciences. Moreover it seems that much of the nostalgia for the purely philosophical approaches – for conceptual analyses – is not based on any principled argument at all but is only a feeling or an attitude engendered by the still recent tradition within the analytic philosophy. I do respect that tradition, at least as much as I am suspicious of its potential (which seems to have been not so impressive). So I propose that the diehards would try to understand the naturalist approach (even though, as it seems, it is becoming a new tradition).

Something must also be said about the specific methodology I use in constructing my scientific neurocognitive framework for the senses and reference of proper names. The considerations are obviously not meant to be extensive. I am not going to give any detailed arguments against any particular models found in philosophy of language and cognitive science; general points will do.

I am not using connectionist models. Fruitful and promising as the approach ("connectionism" or "parallel distributed processing") has proved to be in some areas of cognitive functions, my reasons for excluding it are mainly the following. We are not yet in the position to see clearly which of the connectionist models, and explanatory principles, are the most powerful in the sense that they would correspond to how the actual neural networks in the brain and in the central nervous system represent and process information. The most obvious problem here is that the connectionist networks (even the ones with recurrent architecture) only imitate the gross features of the actual neurons and networks. For instance they lack the essential neuromodulatory effects due to the second messengers, neither are the hormonal effects on the activity states of the networks taken into account. In general the whole synaptic dynamics is ignored: the interactions of transmitters and receptors, the functioning of different types of receptor channels, the role of calcium, and other such factors. Moreover it contributes little to our understanding to say that the information encoded in the networks resides in the weights between the individual units. One can well ask "what is in the number?", i.e. what the numerical value of the connections strength represents? If all we had at our disposal were the connectionist networks, we would be at a loss with respect to what and how the mere connection weights represent and enable the network to use the information residing in it. Neurobiology has,

fortunately, made some theoretical advances about the mechanisms and even of the molecular steps of the impulse transmission and the information preservation. Though that understanding is also quite rudimentary still, it is my I-hope-at-least-somewhat-educated guess that as long as the connectionist models and simulations keep on lumping together all those intricate factors relating to the synaptic processes into numerical values, their utility as explanatory models will come to an end sooner than later.

And not only are the current networks poor with respect to the synaptic dynamics, they are also quite out in the dark with respect to the overall network dynamics; the lack of many relevant functional analogues to the real network dynamics is only too obvious. To mention only a few: the recurrent nature of the networks should be multiplied so that there would be recurrency within recurrency. This means that the amount of the processing units (hidden units) must be "grown". Something like the cortical columnar structures must be added, if they are not emerging from the dynamics of the networks. In short, connectionism is still too coarse-grained approach (even if it seems to be on the right track in light of some of the tasks the networks have been given to solve so far).

My general problem is how the mechanism of reference of proper names works neurally. I try to make way toward the solution of these problems by building a framework that focuses on the neural processing of the cognitive information. So I must use the language and concepts from both neural and cognitive domains. The former language uses such notions as synapses, recurrent connections, neural networks and hemispheric lobes, notions that are explained in due course. I presume that the reader is more familiar with the cognitive language, that he has an adequate grasp of such notions as representations, mental models and intentions. (Not that these notions would be more uncontroversial than the former notions.) But the reason for the use of the two languages is simple: we have to be clear on what the neurons and the networks are processing, for that ultimately constitutes what cognitive activities are going on in our brains.¹

The execution of my project is deliberately somewhat loose. This means that in some places I will be clutching any straw that have proved to be at least relatively robust and quite uniform even if only tentatively accepted. But I will occasionally also go to the other extreme and pick up an idea which seems to me very interesting and worthy of development. To the latter class belong the hypotheses of the re-entrant (recurrent) processing throughout the levels in the neural systems, and the application of that idea with the temporal synchronization of impulses to explain various traditional problems of reference (like Frege's data). It may seem that the occasional clutching is an indication of a theoretical opportunism. But so be it, for I do not sympathize at all with that kind of an approach, usually called normative naturalism, which tries to find and justify some general and necessary aspects of "the" scientific method conducive to truth (or warranted assertibility). Self-correctiveness and some other features are most probably the constitutive features of "the" scientific method. But science is also a historic enterprise and viewed from the neuronaturalistic perspective its methods and tools are our own creations, hence they can also be made better still as the result of our cognitive and practical advancement. Especially in the reference research we are still in the early days, so it would be premature to insist on some normative methodological

rules when we do not even have decently developed cross-disciplinary theories on which to apply the rules.

Still one might ask why it is supposed to be important to construct a neural model. We may even assume that everything that the referential use of language requires takes place solely in our brains, in that the contexts of use are taken into account because they are informationally projected inside the brains and informationally constructed there as representations. And taking for granted by now that my arguments against the explanatory adequacy of the direct-causal theory of reference are correct, why still not just concentrate on the familiar way to do philosophy of language and mind? That is, to theorize only more specifically about the cognitive representations and their dynamics without bringing in the brain? In addition to the pure fun of doing it, there is a simple but compelling answer to my mind: taking the neural facts and hypotheses into account we *constrain* the philosophical and the purely cognitive accounts. For example there are lots of good studies of discourse semantics, referential accessibility and such subjects, which have provided much empirical data. But the problem with them is that because they are not linked to the actual underlying neurocognitive dynamics giving rise to the data, the explanations they provide are bound to remain insufficient. And eventually we will be finding ourselves in the situation in which the only way to proceed is to “go neural”, so why not start now - for we also learn through errors. If my neurocognitive framework turns out not to be on the right track, at least that track will have been trodden and another more fruitful can be tried next.

Neurocognitive information

"Cognitive" and "information"

Let us begin with an explication of the notions of cognitive and information. First of all, "cognitive" may seem to carry more weight than it actually does. It has become an umbrella term under which myriads of more specific hypothetical postulations find their home. To mention a few: scripts, frames, schemas, concepts and percepts. Basically when something is cognitive it is antibehavioristic; "cognitive" is both a descriptive and an explanatory label referring to the internal processing structures and mechanisms which interact with one another and with perceptual information to produce both behavior and new informational and representational states in the nervous system. So in this sense cognitive is thoroughly functional. (Or functionalistic but without any ontological commitments to philosophical functionalism on my part here; nervous system just has the general three-tiered functional architecture of input, internal processing and output.)

What about information? It has been carrying the main explanatory burden in my arguments. "Information" has become another very general term used even as a metaphysical category (as is the case with Salmon's semantic framework where it is something "eternal" (Salmon 1986)). In this study the material base of information is neural processing in all its forms from the biomolecular level to the level of the neural networks of networks. So I shall not deal with the physical information outside an organism or a cognitive agent in any detail; I assume that the Gibsonian view about it is largely correct in regard of the events and objects (Gibson 1966) (see also (Evans 1982: 122-9)).² Obviously the physical informational patterns in the brain contain also the non-perceptual information about distant events and objects. It is understood that

information is transduced into the forms of the neural impulses at the detector sheets that are the boundary between the organism and its environments. I use the expression "neurocognitive information" because it directs our attention to the neural systems involved in the processing of the cognitive information. Cognitive information is usually taken to be conceptual, but not exclusively. That is, cognitive information tells what events and objects are. The referential use of language, *a fortiori* the uses of proper names, utilizes that information because the referential use of language is a form of human action and as such it is intentional because cognitive and representational.³

We are not aware of the amount of the neurocognitive informational goings-on in our brains that lead to the referential uses of linguistic expressions. How then can we find out about the varieties and amounts of that information? We have to be able to do that lest we end up permanently with the folkish analogical and phenomenological approach the neo-Fregeans have dwelled in. It seems to me that to begin with the most revealing source is cognitive neuropsychology. It studies the impairments of the cognitive capacities due to lesions like injuries and strokes ("experiments of nature"). Sometimes the impairments turn out to be quite selective. These cases, painful and confusing as they may be for the victims and their close acquaintances, provide us with ample and sometimes quite bizarre material. The relevance of the cognitive neuropsychology, in comparison to mere cognitive psychology, is that the cognitive functions involved can be related to relatively specific brain areas and to specific subsystems, sometimes even molecular ones (as is the case with the research on hippocampus).

To mention some relevant types of cases, a defect in a certain area of brain may cause, say, an inability to find correct words (naming anomia) but spare the comprehension of the words when one hears them. Clinical record contains patients who are unable to name familiar persons, species of animals, colours, fruits and vegetables, words of abstract things, words of concrete things, and indeed proper names (Semenza & Zettin 1988, 1989). The prosopagnosics, in turn, reveal the importance of faces in communication and in identification of persons. And conversely, the information about the bearer of a face may not be affected, which enables successful referential act taking place. The implication of this is that proper names are not necessary for successful reference. Like face(images) proper names facilitate reference, but if you are a name anomic, usually you still know perfectly well who the person in question is, and you can use that information to achieve a referential act by exploiting it through other means, descriptions most likely. A particular case may be in order to illustrate this. I quote at length for I think that the quotation is self-explanatory.

"A further observation was that RFR, although unable to name people, could discriminate a famous face from an unknown face. He was also able to give the surname of a famous person when cued with the first name plus the initial letter of the surname. This suggested that RFR might know more about people than revealed by his basic naming ability. Subsequent observations indicated that RFR often possessed considerable information about people, but this never extended to any events that these people have featured in... The conversations describing his account of two friends illustrate the point:

'He's a colleague I've known for many years. I think we are both in the...unit.'

'Can you describe him?'

'He is a rather chunky individual. Everything is large except for his height which is about 5'10". I'm not sure if he has ginger hair...He is an outgoing character of Scottish descent. I vaguely remember joking references towards his Scottish ancestry and his love of whisky. He does have a very attractive wife with the very apt first name of Eve.' (Parkin 1996: 210-1; I have omitted the second conversation.)

We could not fail to see that the descriptions the patient gives contain uniquely individuating information - excepting the extreme counterfactual cases of "massive duplication" as irrelevant because as philosophical *Gedankenexperiments* they could not be prevented by any means. (Valuable discussion of the points just made can be found in (Ellis & Young 1988: 5-25) where the authors take up also the methodological and the ontological presuppositions of cognitive neuropsychology.)

From sensory to cognitive information through some functional neuroanatomy

Let us turn to some of the relevant results produced by the functional neuroanatomy and the neuropsychological investigations. The most relevant impairments to my approach are those that have to do with the neurocognitive processing of objects and persons. But a word of warning is in order: when I speak about shapes, colours and such properties as informational, do *not* take it the way that I would be saying that the brain executes its processings *with* such items, and builds up representations from them in a strictly compositional way. Rather all the involved networks process information in a temporally extended interactive manner, in a type of "global" processing within and between the many networks in which the neural representations are constructed.

Object recognition defects show that the different types of information processed involve visually and tactually the shapes, textures, sizes, motions, sounds and colours (not to forget tastes and smells). The information about persons, in addition to those just mentioned, involves such pieces that constitute the ability to recognize faces visually (and tactually by the blind). All these types are sensory information and they are distributed in the different areas in the brain, but they are all processed simultaneously in parallel and in reciprocal fashion. They so to speak come together to form perceptions. If that were not so, you could not see and feel an object at the same instant, or you could not detect movement by auditory means on the base of the Doppler effect. Motor information about movements and action patterns is included also in the informational pool we are interested in here. For example it is known that neurons in specific motor areas fire when one is imagining motor actions while not executing them (Georgopoulos *et al.* 1989), so it is quite possible that elements of motor information get activated when one hears someone mentioning, say, a name of a tool. Information about the functions, and uses, of objects seems to be somewhat distinctive with respect to the information about the shape, texture and colour. Ellis and Young propose that the information about the use is linked to the structural properties of the object rather than to the "semantic properties" (Ellis & Young 1988:

56). This may be so, but I want to suggest that there is another explanation of the fact that some patients can, for example, mime the uses of the objects they could not name (i.e. when the information about the "semantic properties" is severed). The phenomenon may be related to the motor systems: the typical use of an object is coded as motor acts or movement patterns which can be engaged as long as some effective connections are preserved from the "semantic properties" to the motor areas. This could be related also to the now well known fact that severely amnesic patients, who can not remember their post-lesional experiences, can learn with repeated practice some manipulative tasks (like the Tower of Hanoi problem) even when they do not recall ever being confronted with such tasks before when queried at any instant after the practice. But be that as it may, my point is that it is rather artificial to separate "semantic" neurocognitive information from the motor information and information about the functions of objects (and typical actions to be executed in normal social events, like visiting relatives - especially boring ones).

In addition to the point about the motor information I want to propose that the relevance of emotional and affective information should not be forgotten either. If one is after semantics that will be able to provide scientifically grounded explanations of behavior, the referential uses of language included, then perceptual, cognitive, motor and emotive (plus even endocrinological) information should all be eventually incorporated into that framework. The reason is that they are all functionally intertwined in the brain. From the perspective of the functional neuroanatomy this interrelatedness is almost a truism, because the regions involved in the processing of these kinds of information are connected to one another, both directly and indirectly, via interposed networks. For an explicit suggestion, on the ground of the growth of the neural connections from the sensory areas to the frontal lobe areas, that proposes that *object identification* and *significance* is tied to emotive processing, see (Pandya & Yeterian 1990). (On the psychology and neurobiology of the general issues, see (Buck 1987) and (LeDoux 1987).)

As a token of empirical justification of the general perspective let us take a look at Capgras syndrome. It is a remarkable phenomenon in which a patient thinks that for example his spouse is an impostor, not the "original" loved one. The most favoured current hypothesis is, roughly, that the connections between the limbic system and the visual areas are severed to the extent that the patient has lost his feeling of familiarity with respect to his spouse (Young 1998). This causes a delusional misidentification. This may show up in the referential practice of the patient. One is, one thinks, not referring to one's original loved one but to a sort of *Doppelgänger*, an impostor. It is known that on some occasions this kind of delusion has led a patient to commit the murder of the "impostor". It seems quite clear, indeed, that the syndrome involves a disruption of the emotional evaluation of the perceptual and cognitive information related to a person. Note also that Capgras syndrome is not restricted to persons: "One may extend this argument to other reduplicative paramnesias. Places, objects and so on are not affectively neutral and so the absence of an emotionally charged input would produce the feeling of recognition, but it not being quite right." (Ellis & Young 1998: 236) With respect to the reference of proper names this is all for the good, if I am allowed to say so in this connection, because places and even some objects are given proper names.

Another type of clinical case in which the descriptions do not "reach" the emotional elements of the senses related to persons is the following (Tranel & Damasio 1993). A patient who is severely amnesic and unable to form memories of his encounters exhibits differential covert responses towards the individuals who behave in different manners towards him. He tends to become uneasy and to avoid the company of a person who has been unfriendly. The contrary behavioral pattern is evident towards a person who has been friendly and helpful. And a neutral attitude is adopted when a person has been neutral to him. This case is grateful because the cognitive information is minimized: because the patient is amnesic he does not form consciously accessible memories, and so could not give identifying descriptions of the persons (for example by saying "that darn ginger-hair-pieced bloke who always makes me angry"). But it seems evident that emotional information has obviously become registered and has had its effects on the differential patterns of behavior towards persons and on the valuation of persons by non-verbal means.

The motor and visceral information processing do not show up explicitly in the surface descriptions (except, maybe, in rare cases). But that could not be used as an argument against the enlarged view on semantics that incorporates a version of the Fregean sense theory of reference. The descriptions are the temporal publicly expressed products of the sea of processing in the brain and the whole nervous system (both central and autonomous). What counts in this framework are the varieties of the internal information processings leading to referential uses of language. With proper names this presents one more step away from the underlying informational dynamics because the descriptions are so to speak passed by. But if asked about the referent of a proper name, the focus turns on it, i.e. it becomes again the expressed end product of the processing. (The main "pointer" in the Recanatian encyclopedia entries.)

However one may object that this mesh of the kinds of information is ultimately un-Fregean. For did not Frege deny categorically that *Vorstellungen*, i.e. sense perceptions, images, feelings and such *subjective* things do not belong to the domain of semantics and cognitive significance. *Vorstellungen* are intersubjectively non-sharable occurrences, therefore not amenable to thoughts that are shareable entities, hence objective.

It seems to me that the answer to this *prima facie* pressing counter is the following. I have already motivated my neurocognitive approach by naturalism (though somewhat fleetingly so far). This means, minimally, that any "third realm" postulations are strictly unacceptable. The positive side of this naturalism is that thoughts are located in the brains of the speakers, together with *Vorstellungen*. For I am in search for a modern scientific version of the sense theory of reference, and because the modern ontology of psychology and semantics is unashamedly materialistic, we must find the neural reconstructions of the allegedly abstract entities in Frege's semantics.

What seems to be a more important consequence is that *thoughts become also subjective in this approach*. But it does *not* follow that they become *wholly* subjective. To begin with it should be realized that *Vorstellungen* are not wholly subjective either. To suppose that they are was Frege's mistake. The argument for this should proceed by each instance of the types of *Vorstellungen*, by the different perceptual modalities, emotive attitudes and so on, but here we have to be content with only a general argument. It can be admitted that *Vorstellungen* are subjective and occurrent events.

But this by itself does not preclude sufficiently objective features from being elements of them. For instance when two persons perceive a tree up on the hill, their perceptions differ from each other because the angles of sight differ respectively. However both perceptions are objective because they are of the same tree. Were there any doubts, they could approach the tree and meet each other by it. So we see that in fact *Vorstellungen* do not differ qualitatively from more cognitive acts and their contents. For after all, as I have emphasized on a couple of occasions before, Frege allowed that different persons have distinct senses related to the same referent of a proper name, distinct in that they are not grasped by both (though they are graspable by both). This means that their respective senses are not the same, but that does not hinder sufficiently objective communication and mutual comprehension when the referent is the same, as Frege pointed out.

But what about those *Vorstellungen* which are directed internally and without an external object? Are not feelings of, say, anxiety only subjective in the sense Frege meant in that they have objects, viz. the feelings themselves, which are not graspable by others? For how could anyone grasp another one's inner feelings? But even here I fail to see that the extent of the subjectivity of these cases forecloses them from having sufficient degree of mutual comprehensivity resulting in objectivity. The reason is rather obvious when one becomes to think of it. It is arguable, and with considerable scientific support, that for example emotions and affects are evolutionary products. They contribute to the well-being and survival of organisms, at least the "higher" ones. (But I surmise that when one accepts biophysiological perspective that is seen to be the case with amoebas as well; their emotions are only rather simple, as when the detection of some poisonous ingredient in the surrounding water causes internal reactions conducive to protection.) Emotions, emotive and affectual information, guide our behaviors toward our surroundings at large, and towards or away from particular types of events, objects and persons in them. They contribute to the well functioning and cohesiveness, as well as to the grounds of joint actions, of groups of peoples. Consequently it is very likely that the kinds of emotions like anxiety are intersubjective and sufficiently similar biologically. This is sufficient objectivity for the semantical purposes.⁴

Frege held that the senses and the thoughts make up the common stock parts of which all competent speakers of a language grasp, this ensuring objectivity of communication. Similarly we can state that the particular kinds of underlying emotional states (anxiety, fear, happiness, and such) form a common stock for all humans to grasp, to experience, so that we can objectively comprehend one another's states of mind and moods.^{5,6}

I think that the cognitive terrain is sufficiently charted by now. So it is time to turn to the search for the neural concepts and ideas by which we can begin to provide adequate explanations of reference. A motivational reminder: this approach stems from the fact, unsatisfactory to my mind that the cognitive models, as well as the neo-Fregean ones, are built on only unbiological, folk psychological and other such coarse and molar concepts. Still, as pointed out above, that does not mean that we should adopt exclusively the "neuron talk". Neurons work in concert, so the new neurobiological approaches should focus on the network dynamics, and develop new images, concepts and structural notions appropriate to these levels of investigation.

Recurrent/re-entrant networks

Neuroscientific background

I mentioned in passing above that one of the most important questions in the current neuroscience is how the perceptions and cognitions emerge to form coherent and unified patterns of representations. Why is it that we perceive unified scenes instead of fragmented, when the various elements of the information processed lie in distinct areas in the brain? This is called the *binding problem*. Or, rather, there are *several* binding problems to be explained. For although the perceptually unified patterns that we studied are mostly visual, no less remarkable are the coherent patterns of the other sensory modalities *and* the cognitive patterns we call thoughts and concepts. Nor should it escape our notice that the motor actions and the movements we execute are harmonious and "adequate-to-the-aim", not jerks and discontinuous (as they are in some severe cases of lesions like dyskinesias). And the emotional patterns like affects seem to conform to the binding pattern also. This, at least, is what I propose: bindings occur not just within perceptual modalities but within the "conceptual modalities" also, as well as within the motor and emotional ones. And even this generalizes: the most remarkable binding problem of all (should I say "super binding problem") is how perceptions, cognitions, emotions and motor processes form unified and coherent overall patterns so that we are able to adjust ourselves in the changing environments to survive and flourish in appropriate ways. "Binding" is a good label in this connection even if perforce somewhat metaphorical. "Construction problem" or "emergence problem" might be less metaphorical, but "binding" is what has bound researchers to itself. The binding problem gets its rationale from the empirical observations that there are no "grandmother cells" or "unique central integration areas" in the brain (no neurofunctional or anatomical Cartesian Theatres). True, there are integration areas in the brain in that large neuronal groups and columns processing similar information are clustered together or in one another's vicinity. This is evident from the processing properties of the neurons for instance in the primary sensory areas. These sites are required in the neural information processing streams for perceptions to emerge, but their activity alone does not suffice for perception. Other areas and the neuronal networks linked to them, and the modules processing much more generalized and abstracted information (conceptual/cognitive information), must be automatically recruited to the processing stream giving rise to the overall activity patterns, to cognitions. So how is it that from the sensory stimuli the brain constructs coherent and unified representations by binding? Why is it that we see a whole scene, a whole apple, a face, and not mere patches of colour? How we grasp concepts like "chess", "inquisition" and "bed"? That is, how the neurocognitive information residing in the networks gets together in a coherent manner, relevant to the task? In particular how is it that when we are talking about a particular person, referring to her, the bits and pieces of information get activated and get together to tell us that it is just that person, not someone else?

During the last twenty years one kind of general neural mechanism proposed to explain the binding problem have gained ground formidably. I will elaborate on this mechanism soon, but before that we need to present another seminal idea in the recent theoretical neuroscience. The idea is that neurons and larger networks consisting of them form reciprocally interacting groups, *re-entrant* or *recurrent* networks. The patterns of the interactions are dynamical and flexibly changing, which means that the

strengths and combinations of the transmitted neural impulses change: sometimes some connections remain silent, though they are still functional; new connections are formed and others die because prolonged lack of use, lesion or natural cell death. This recurrent pattern is hypothesized to pervade through every functional and anatomical level in the brain, from the connections between the individual neurons to the microcolumns through the macrocolumns to the connections between the specific areas within a lobe and from them to the massive interlobe connections and finally to the hemispherical connections through *corpus callosum*.

The synaptic connections of the re-entrant networks form also diverse structural patterns. They can be divergent and convergent to different extents; the axon collaterals can be arborized to varying degrees; the connections may be layer-specific or diffuse. (The cortex has 6 main layers with different and quite celltype-specific input-output properties.) There is variability at the biochemical level also. It is typical that the action potentials have different pre and postsynaptic chains of effects with temporal continuum from microseconds to days, and longer. The amounts of various transmitters released relate to the received impulse frequencies and cause the second messenger cascades inside the postsynaptic neurons, which in their turn affect the gene expression, for example through the CREB family of proteins, leading to morphological changes in synapses and transmitter dynamics all over again. The current view seems to be that these structural changes contribute to, or even sustain, the long term memories in the cortex.⁷

The other side of the neural coin is the relative morphological stability, or robustness, of the basic processing structures (excepting the growth and death of synaptic connections and the enlargements of the areas due to the former mechanisms. Much of this stability is governed by genes and their expression during ontogenesis (though it is now known that experiences can change them to some extent). In part this means that the re-entrant projections by the axon collaterals are quite specific and localized in the receiving end, for example they are tied to specific cell types and connection layers. The generalization that can be drawn from the current state of knowledge is that the dynamic processing patterns within the structured neural pathways range from stable responses to properties in the environments to even chaotic responses of the networks, settling temporarily down to "local minima" or approaching the "basins of attractor". (On these chaotic patterns in the brain, see for example (Freeman 1999)).

In this connection let us take a quick look at the functional morphology to give us a conception of the levels of processing in the brain. The microcolumns are about 30-50 micrometers in width and contain around 100 neurons. Their connections are both vertical and horizontal, both excitatory (mainly pyramidal cells) and inhibitory (mainly basket and double bouquet cells). The inhibitory properties are prominent in forming the neural clusters that will form the columns in creating functional borders with respect to the specific information processed in the columns. The columns are made up of microcolumns, from 100 up to 300 in number. The width of the columns is 300-500 micrometers in diameter. Their primary pattern of constitution is modality-specific and receptive field-specific. For example fast and slowly adapting receptors in the skin are segregated within the same neuronal receptive fields in the cortex, i.e. areas in the skin that are projected to the neurons in the somatosensory cortex. From these columns is constructed the "somatosensory *homunculus*", the cortical representation of the receptor area layout of human body. The same principle applies in all sensory systems

with respect to all sensory features processed. But columns are also found in other cortical areas. Their processing specificities are more general and abstract features (see for instance the "shape columns" in the inferior temporal area in (Mountcastle 1998: 185-9)).

Columns form macrocolumns (or "hypercolumns") that process different features of the modality or task-specific information like that needed for face recognition. The primary and the secondary sensory cortices are formed from these different cortical areas and regions. The cortical lobes are formed in turn by these informationally connected areas. They can also be characterized by the types of information and by the tasks they execute. Generally characterized the functions of the frontal lobe relate to movements, motor actions, planning of those actions, features of short-term memory, and to personality traits (in this case the involved areas are the prefrontal ones with connections to the limbic system nuclei). The temporal lobe is primarily engaged in processing auditory information, variety of memory representations and language functions. The parietal lobe processes somatosensory and spatial information (related to movements and actions); and the occipital lobe performs visual tasks. Summarily, then, the re-entrant connections run both within all neural assemblies and between them from the "higher" regions and assemblies to the "lower" ones, and back again.

Edelman's TNGS

For a quite long time we were without a general but sufficiently detailed theory of the ways the brain represents and processes information. Currently this lack has become remedied to a reasonable extent. In the last twenty years there have been several attempts at such frameworks. The one that seems to me to be the most promising is Gerald Edelman's TNGS (Theory of Neuronal Group Selection). It is a framework to explain consciousness by neurocognitive structures and dynamics. But here I adopt for my purposes its basic machinery for it can be used to explain the reference of proper names (and, by a rather obvious implication, reference in general) - or so I will try to show in the following. Consequently I will discuss TNGS quite extensively because that is indispensable to get an adequate picture of the neurocognitive approach I am arguing for.

The theory is based on the population properties of the neurons and the neural networks. The core idea is analogous to Darwinian evolutionary theory: selection upon variation (Edelman 1987, 1989).⁸ One supposition of this theory is that *no* stimuli from both the external environment and the organism's internal milieu are *labeled* or *categorized* in advance when they come into the neural processing streams. This is contrary both to the practice and some deep-seated assumptions in cognitive science and classical AI. Organisms with sufficiently rich nervous systems construct the to-be relatively stable categories from the input through the internal processing and motor actions, as well as the results of the latter through external feedback. The results of these processings are constantly fed back to the wide variety of networks in the nervous system. So categorization of the information is generated by the selection upon this neuronal variation of and among the neuronal groups. This means that the nervous systems i) have to develop in an adaptive fashion through the encounters with their surroundings, and ii) being modulated by the internal informational messages. Without going into the details it suffices for my purposes to describe shortly the three selectionist mechanisms operative in the brain (Edelman 1989: 43-9).

The first is a developmental one during which primary connections between the neurons and the neuronal groups are established by genetically and molecularly guided cell migration, proliferation, cell death and functional synapse formation. Stimuli already drive this stage, for it is known that properly functioning networks emerge from the interactive but competitive stimulus processing. This overall development leads to the *primary repertoires*: these are the variant neuronal groups within the anatomical areas.

The second selectional stage, overlapping with the first, is an experiential one during which the neuronal populations of the synapses are strengthened and/or weakened to various degrees. It seems that these patterns of activity are most likely responsible for the emergence and the stabilization of both the microcolumns and the columns by excitatory and surrounding inhibitory connections in the primary repertoires. Note that it is still not the stimuli as labeled, as about something outside the organism, that drives this process. The signal correlations are statistical, i.e. the establishment and the stabilization of the connections is more a mass effect than stimulus specific. The main result of this stage is the formation of functional connections between the different and distinct areas in the brain.

Only in the third phase the specific informational features of the stimuli become central to the processing. This takes place in what Edelman calls "*re-entrant mappings*" (or equivalently: recurrent mappings) that serve the adaptiveness of the organism to its environments. The re-entrant mappings create the primary sensory and motor representations, or maps, and their mutual interactions. More specifically, this is stimulus exchange by re-entry through

"...temporally ongoing parallel signalling between separate maps along ordered anatomical connections. Reentrant signalling can take place via reciprocal connections between maps...(as seen in corticocortical, corticothalamic, and thalamocortical radiations); it can also recur via more complex arrangements such as connections among cortex, basal ganglia, and cerebellum." (Edelman 1989: 49)

Thus the features and scenes are constructed by the mappings in the visual cortex (and connected subcortical nuclei), auditory representations by the mappings in auditory cortex (together with some subcortical nuclei), and so on. Once the organism has reached this stage in its development, the modules that process sensory and motor information are fully functional and the organism is capable of achieving perceptual categorizations of both the world outside and of some of its own states. There is nothing linguistic in this yet; in humans language develops upon these perceptuo-motor categorizations, and the recategorizations of them.

The re-entrant mappings are many-many. Distinct neuronal groups within one area take part in the processing of the same occurring stimulus features, and one group with dynamically varying processing properties (like those in the levels of the synaptic activity) takes part in many different but similar stimulus features.⁹ For instance there are many groups that process motion information, and there are groups that process, say, orientation and colour information in addition. The same is true of the higher level networks, those further up in the processing stream in the "association (tertiary) areas".

These re-entrant many-many mappings Edelman calls "classification couples" and "classification n-tuples". Those are formed through the selectional mechanism in which one neuronal group in a map is correlated to another group in another map. For example a group of visual feature detectors for some specific shapes in a map is related to neurons responding to light touches, forming another map through the selective re-entrant connections between them. The classifications are not restricted to sensory maps but involve motor maps also. Later in the cognitive development of an individual the mapping of these maps themselves leads to categorization of the information in and between the maps as well as the classifications they embody.

To achieve perceptual categorizations from the re-entrant mappings the formation of *global mappings* of classification couples and n-tuples is required. These create the specific spatiotemporal continuity of the representations of external objects and events. A global mapping is essentially what the label says: a dynamic structure engaging many brain areas within which sensomotor re-entrant mappings interact with other regions, but mainly with cerebellum, basal ganglia and hippocampus. The latter are the areas that, according to Edelman, provide for the (different kinds of) temporal sequential orderings of the representations in the global mappings. The processing in these global recurrent systems underlies the categorizations built on the re-entrant classification mappings, and even further categorizations and generalizations of them. These latter patterns are *concepts* in Edelman's framework. Hence in addition to the object, action and event representations (categorizations), the relations between these become comprehensible. For instance causality is one such categorization from objects and events and their mutual relations from rudimentary contacts to the folk concept of causality. (The global mappings have also reciprocal connections to the nuclei in the limbic system and to the brain stem nuclei, which contribute essentially to keeping the organism alive and aware.) In short, global mappings equip the organism with spatiotemporally continuous representations of objects, actions, events and the relationships between them. The connections to the limbic and brain stem areas provide the continuous updating of the value of the categorizations for the organism in the sense that the evaluations guide its behavior with respect to its surroundings. Such evaluations are not conscious in many cases but "visceral" and neurophysiological in that they alarm the organism or direct it towards or away some object or action. The limbic system is central in these evaluations.

In this connection I could not resist offering a hypothesis, albeit more on the speculative side than (as of yet) based on empirical indications. What if we could have a middle ground between totally selected-for neural groups and the dying-out of the unused synapses (both during the neuron deaths in early infancy and at later stages in life)? That is, what if *in addition* to the gene expression induced structural changes in the neurons and synapses there are "spare synapses" ("sleepy synapses") and/or connections? These would be in silent state – except that to keep them alive and healthy they generate spontaneous spikes. In this way it might be possible to explain for example the quite quick "one shot" consolidations of memory representations. (The gene-induced changes just seem to take much too long time for that.) In general my suggestion is, then, that there is some amount of reserve connections ready to be activated when the storage and/or functional capacities of the selected networks run out or are unable to function for some neurobiological reason. This could also account for the fact that the selectional mechanisms allow for the plasticity in terms of the

number of the connections, and not just in terms of the modulatory or spike frequency, as well as strength effects, among the already selected neuronal groups.

But let us go back to the global mappings. An example helps to comprehend the nature of the global mappings. Hearing someone talking about the players of the team Brazil activates many distinct but informationally related columns that process the properties in the auditory stimuli as words (in the temporal lobe). From these the activation spreads to one's memory items, to the related global mappings or parts thereof, mainly in the inferior temporal and occipito-temporal sites. As a result the hearer entertains images of the individual players, maybe also memory flashes of Brazil's games during the World Cup, of the technical virtuosity of the individual players and such pieces of stored information. These all come together - the meaningfulness of the representations emerges - from the categorization and the individuation of these elements through the spread of the processing, up to the association areas both in the occipito-temporal and inferotemporal regions, and also in the parietal areas (especially if one has experiences of playing football). Related experiential information in the premotor and supplementary motor areas residing in the frontal lobe may also be activated and adding to the somatosensory activation. (The frontal engagement may also express cognitive information retrieval (Sergent *et al.* 1994).) If the whole process ends up with a pointing to a particular player - from the information of blue shorts and yellow shirt with number 9 on its back and the player pressing forward between the defenders - and the speaker uttering "Ronaldo is in top shape again", then the involved areas include also the posterior superior temporal gyrus (Brodmann's area 22) and inferior frontal gyrus (area 45), for they are involved with name-familiarity and name-profession tasks. The neural representation of the proper name resides most likely in the left temporal pole (area 38). It is to be noted that the usual feeling of familiarity, occurring just before the "popping-out" of the proper name in question, may have much to do with the limbic system, and it seems that the emotional colouring comes from the subcortical nuclei of this system, especially if one happen to be a football fan. This overall processing takes place within 500 milliseconds but the time window is taken to involve long-term memory item activations and non-conscious elaborations of it. This is quite a long span in neural terms since there are indications that the "pure" recognition of an object, say a face, may take only 20-30 milliseconds (Rolls & Tovée 1994).

If I have understood it correctly this whole processual engagement can form assemblies of global mappings, widely distributed and recurrent processing streams that the stimuli have prompted. It could also be that when the global mappings are processed in concert a couple of times, a more general, or wider, global mapping may emerge as a relatively stabilized entity: one's *concept* of, say, Ronaldo. The example should make it plain that we go through these kinds of engagements of the global mappings all the time. They steer us cognitively (including movement patterns and emotions). Because of the recurrency between the networks, the informational patterns generated in the association areas have their say back on the processings going on in the sensory areas. They guide that processing by making the context meaningful (as the football event in the example). It must be emphasized that all this takes place almost simultaneously and in parallel within the 500 milliseconds range, *not* as some kind of "interpretation" put on distinct sensory materials. The overall pattern of the dynamics within the neural architecture indicates that this rather common interpretation account is neurofunctionally incorrect. Another corollary is that those

theories that see perception, though correctly conceptually molded, based on "unconscious inferences" might not be correct either – at least not with respect to the rather plain talk of “inferences”. I propose that the activation patterns the global mappings embody are not inferential in any extended sense of that notion. Rather they might function as *informational coherence detectors*. By this I mean that because of earlier experiences the stimuli spread through the different networks within the global mappings informationally engage those parts of the global mappings that sustain the coherent and appropriate-for-the-task representations. It is more our persistent habit to talk about inferences than what is really going on at the neural level.

Third corollary, and the most important from my view, is that the global mappings manifest *activation of concepts*, including *individual concepts* of persons and objects, as the bearers of proper names, connected to the global mappings via the neural representations of the former. These individual concepts are not linguistic in the sense that the multitude of the different kinds of neurocognitive information that the global mappings engage between the maps, and the specific pieces of information residing in them, are not linguistically wholly articulatable. But nevertheless there are no serious obstacles to taking the selectional informational connections that form the re-entrant categorizations to *be* individual concepts. The linguistic expressions that are produced as the results of these particular (and context dependent) categorizations through the global mappings are very compact reports of the informational traffic in the global mappings. Consequently the senses of the linguistic expressions, *a fortiori* proper names, are not primarily linguistic - let alone that the Fregean neural senses could be identified with linguistic meanings.

Neurocognitive information in the global mappings is engaged into the processing stream because it resides in the neuronal groups and is linked and functionally correlated to other neuronal groups that are, even if only transiently, selected as active due to the similar informational features encoded in them. It seems to me that this is the most likely dynamic-cum-structural explanation of why not just the perceptual but also the conceptual aspects of cognition form coherent patterns, and not just patchworks of representations without meaningfulness. Much of the coherence, and unifiedness, is explained by the topographic feature organization in many areas. I would not be surprised if such feature clusterings will be found uniformly in the association areas also (though they could be in the form of macrocolumns or some such higher order processing configurations). The activations of the various portions of the global mappings, leading to the associations between the features in those mappings, bring with them the power to generalize from experiences. The global mappings form also the basis of memory, due to the long-term changes in the secondary synaptic connections (as Edelman calls them). That is, the synaptic changes within and between the synapses of the neuronal groups (maps) initiated by hippocampus result in memories. Most likely the long-term memory storages reside in the cortical secondary and tertiary association areas. Because of the on-line activation and forming of new synaptic connection strengths and patterns from the portions of the previously activated global mappings, memory is not replicative display but constructive process. The philosophical relevance of this is that the contents of the concepts, the individual concepts included, do not remain the same even intrasubjectively (at least for extended temporal intervals).

As described above language does not play any central role yet in the beginning of the categorizations and forming of the memories and the concepts. On the contrary, the development and the proper functioning of all these cognitive functions have to be more or less in place before language learning can begin. The following is a very short description of Edelman's story of language development (Edelman 1989: 147-8, 176-82). Language begins to develop only when *presyntax* develops. Presyntax is the sequential organization of action, in the sense of ordering, in relation to the sensorimotor re-entrant mappings and global mappings. It engages also subcortical structures like the basal ganglia and hippocampus to create spatiotemporally continuous representations of objects, actions and events, and their mutual relations. When phonetic items appear gradually due to the development of supralaryngeal structures, together with the development and the maturing of Broca's and Wernicke's areas, they become mapped in an orderly way to the previously formed action-object and event categorizations. This leads to more constrained patterns in speech through the recurrent global activity, and eventually to syntax proper. The appearance of the long-term memory, especially in the temporal lobe areas due to the continuous dynamic activity of many global mappings via hippocampus, contributes also enormously to this developmental process.

Once language has reached the stage where words for concepts are being linked regularly to each other, i.e. when the birth of the symbolic function takes place, linguistic expressions begin to accelerate the combinations of the conceptual resources. Language makes concepts richer and sharpens them through the re-entrant mapping combinations that the neural representations of the words connect and facilitate, as portions of them but also as wholes. Linguistic expressions are the products of those mapping processes that engage speech schemas, but the expressions themselves also initiate those processes in their turn. So the overall pattern is parallel and simultaneous reciprocal many-level interaction.

I want to propose that this is the stage when an infant has received all he needs to master the referential use of language. The neurocognitive information he has in his command is in the form of sufficiently rich and refined conceptual structures (though still somewhat rudimentary compared to the structures of the adults). Beginning from the object-action mappings an infant has advanced a long way and has acquired term-object, term-action, term-action + object schemas. All of these are capable of creating new informational assemblages, some of which are transient and some of which become part of the memory stores. At this stage the infant never looks back. In fact he could not do so, for there is continuous informational traffic going on in the brain.

Now I propose that the basic schema of reference I postulated earlier, when I criticized the intuitive arguments of the direct-causal theorists, is fully functional at this stage of the language development. The child has acquired the mastery of the referential use of language, not restricted to proper names but involving demonstratives, indexicals, pronouns, and referential uses of descriptions. The basic schema engages the widely distributed informational-rich global mappings together with the developed syntactic abilities that are based on the action-object patterns. Infants acquire perceptual and motor patterns of, and related to, objects (but not necessarily always functionally interrelated (Spelke *et al.* 1995)). It seems natural that these kinds of patterns later assume also a role in the term-object mappings. The similarity to Edelman's view is clear: sensorimotor capacities which form the prelinguistic conceptual core and

presyntax are followed by clumsy tries at speaking, which gradually develop to the stage of syntactic patterns. When the symbolic function dawns to the child, the relationships between the conceptual information and the linguistic expressions burst and the entrance to the linguistic community is more or less achieved.

Schemas

I have mentioned schemas in addition to the basic schema of reference. Unlike many models in the mainstream cognitive science schemas are not characterized by being static. That makes schemas more fruitful constructs to be used in the neuroscientifically-oriented studies of the various forms of cognition, including the referential use of language. Because that is especially a-very-good-thing from my perspective let us see how that fruitfulness begins to show. But first, let me shortly review the typical properties of the schemas.

Schemas are usually postulated to explain all the main facets of cognition: in perception, thinking and action. The schemas are taken to represent objects as well as actions and events. In the typical accounts these functions are merged: schemas are postulated as cognitively, or conceptually guided sensomotor items. Generalizing a bit, the schemas are perceptual and memory-based action guiding constructs. They are present already in rudimentary forms in the infants' graspings and sucklings, and can attain very sophisticated forms in adults, as in the finest sport and intellectual achievements (among others).

Most schemas have other schemas as their parts. This constituency is not of static nature but flexible in that it depends on the particular occasion in which the schemas and the subschemas are activated. When you look at, say, a typical kitchen, the structure of schemas-within-schemas displays itself readily in action. You see everything from pots and pans to refrigerator and spices, all represented by schemas within the (typical) kitchen schema. And the kitchen schema may be, on that occasion, a part of a house or a family or a cooking schema. When one begins to show one's culinary aptitudes, more richer perceptual and sensomotor schemas are activated, as well as new ones getting constructed along if the occasion differs sufficiently from the earlier culinary adventures.

The sensomotor properties of the schemas lead us to put emphasis on their dynamical nature. Ulric Neisser speaks frequently of the cycles of schemas in perception in his book *Cognition and Reality* (Neisser 1976), but it seems to me obvious that the motor schemas are also involved. In a similar manner Michael Arbib and Mary Hesse emphasize both the dynamical nature of the schemas and their abundant schemas-within-schemas structures:

"We stress that a schema is both a process and a representation. The formation and updating of the internal representation, a schema assemblage, are viewed as a distributed process, involving the concurrent activity of all those schema instantiations that receive appropriately patterned input...We emphasize that the current assemblage of active schemas may contain hundred of schema instantiations, and the entire stock of knowledge of an individual may reside in a vast network of hundreds of thousands of constituent schemas." (Arbib & Hesse 1986: 54)

Apart from proposing any numbers (for hundred of thousands may be an understatement!) I endorse fully what is said in the above quotation, especially about the processual nature of the schemas. As an abstraction a schema can be considered as a snapshot but the recurrent dynamical properties of the neural networks make it rather likely that the snapshot view does not correspond to any sound explanatory notion. Once the particular cycles of the schemas have begun their run in the specific situations, which can also be internal as when one is going down a memory lane, they remain in active states due to the recurrent dynamics – or so I propose. In this way they also activate expectations that relate closely to the contents of the already active schemas; this is the fruit from the shrub of the informational linkage in the global mappings. The expectations are presumably other schemas, involved in the processing of one's drives and needs and aims (in the limbic and frontal areas, cortical and subcortical sites like amygdala).¹⁰

The expectation schemas relate to plans of all kinds. "Little" plans such as those leading to immediate motor acts, resulting from sensomotor schemas, are common. "Big" plans are multi-layered ones ranging from one's aims today to career expectations and such things. In general we see how flexibly the schemas link together from the minutest perceptions via the meaningful features of one's surroundings to one's varying actions in the temporally extended "action chain" that is supposed to fulfill one's grander aims. This flexibility is neurally manifested in the variable but relatively overarching engagement of all the main cortical regions with subcortical ones.

For the various reasons mentioned above I take it that the notion of schema is the most fruitful postulation with which to explain perception, thinking and behavior in a unifying manner. I emphasized above the non-committalness as to the possible neural realization of the schemas. Relying in this "open texture" of the schemas I propose a two-fold *identification*. First concepts are identified with schemas and then schemas are identified with certain neural processes: the global mappings. The main evidence for these identifications is, first, that there are many and diverse perceptual schemas of objects, events and persons. This corresponds to the construction of the object, event and person representations in the global mappings starting from the primary sensory receiving areas and continuing through the secondary regions to the activity in the association areas. Secondly, some of those schemas involve also motor information (in addition to proprioceptive and kinesthetic information). The global mappings engage the motor areas as was briefly described above. That concepts are hypothesized to involve motor information is unusual, but not unreasonable. To think otherwise is just to clutch to the by now rather old-fashioned conception of concepts as only cognitive entities. Thirdly, the schemas contribute to the planning of action and long-term aims. In the global mappings this role is assumed by the higher frontal areas, mainly those in the prefrontal ones. Especially this shows up in the short-term tasks when something has to be kept in memory for successful action to take place. One relevant task of this type is keeping information about someone in mind in order to refer to him or her, and to understand that what is said is about him or her. So I propose that the engaged schemas in these situations are (parts of) the individual concepts of the referents. (The schemas are most likely only partially activated, depending on the informational impact of the situation.)¹¹

Edelman restricts concepts to portions of the global mappings. According to him concepts arise from the discriminations, recombinations and (re)categorizations of parts of the global mappings, like those distinguishing objects from movements in a mapping (Edelman 1989: 144). Note that because portions of the global mappings can be recategorized again in the formation of concepts, these basic mechanisms are in place when language learning begins, so that through the emergence of the symbolic function the conceptual machinery explodes only in quantity. I concur with all this except that I would not restrict the concepts as emerging *only* from the manipulation of the portions of the global mappings by other mappings. These mechanisms may be what *prototypes* are, but full-blown concepts require more. This presupposes that concepts are perceptuo-cognitive skills, or that they form at least parts of the processual phases of those skills. As such they are tailor-made to be identified with the global mappings. But this is a healthy presupposition because the views according to which concepts express properties and relations, or are individuated by some necessary and sufficient criteria, are currently, and deservedly, in disrepute.¹² I think that the manipulations of the parts of the global mappings is a later stage in concept formation, the stage in which more *refined* concepts through conceptual relations and conceptual hierarchies emerge. Through learning the relations between the concepts, their hierarchies and interdependencies grow remarkably.

However, and in tune with what I have earlier proposed, I would like to go a little further in explicating the schemas as global mappings. In doing so I also take the liberty to enlarge Edelman's theory a bit. The schemas have been typically taken as purely cognitive or sensorimotor constructs, but as I argued there is no principled reason why the emotional and affective features of cognition should be kept out. The type of evidence that I gave shows how intimately those features are tied with the perceptual and "intellectual" ones. As a short rehearsal, think about thinking about someone. Your knowledge of his or her is usually emotionally coloured (to varying extents), depending on who that person is, i.e. what you know about his or her doings (especially in relation to yourself), character, and such things.¹³ Your schema assemblage of the person in question includes among its subschemas also the ones containing emotional information and schemas which dispose you to bring about appropriate, or inappropriate, actions were you to meet that person. This is easy to explain on the level of the global mappings. In particular the limbic structures and nuclei contribute their specific effects on the cortical processes (and *vice versa*) and this brings the emotional ingredients into the whole processing stream of the global mapping, the schema assemblage of that person. Consequently I propose that the global mappings can include the engagement of the limbic functions as well, and especially because they partake in directing one's thinking and action by the emotional and "visceral" evaluations of objects and events.

Another important property of the global mappings is that they enable us to recognize individuals. This is mostly due to the perceptual information they embody. As the neo-Fregeans, notably Evans (Evans 1982: ch. 8), have argued the ability to recognize the referent, hence also to reidentify it, is a very important cognitive function. This is especially so in those cases when one could not characterize the referent by any identifying descriptive means, as the case is often with faces. The global mappings explain this feature of the (re)identifying recognition because the information residing in them can not be articulated without a remainder, sometimes a considerable one at that. But it seems to me that not only the linguistically inarticulable perceptual

information helps in recognition but emotional information is able to do that also. For it sometimes happens that one has a feeling which he can not put into words, a feeling which tells one whom is the object of one's experience. For example one might vaguely hear a voice and/or noises that have a particular "flash" and that tells one who is responsible of those voices and noises. A case in point is an event where one is among other boring people and suddenly gets a feeling of warmth and joy upon detecting noises made by someone approaching, that someone being your loved one. In this case there is the interplay with the perceptual and the emotional information but it is the latter which enables one to identify the referent: very similar noises, not consciously distinguishable, could have been made by some other person without that engaging any emotional reaction. On the basis of this example I dare to propose that emotional information is sometimes even capable of executing referential function (though of course in relation to the more cognitive senses one possesses). The importance of this possibility is that the description theory could not account for it, but our (neuronaturalized) Fregean sense theory explains it: some global mappings have networks in some limbic nuclei as parts of them, and these contain, or can access information, that enables one to single out the referent. In short, and contrary to the widely held opinion, emotions and moods can be intentional and even referential.

Individual concepts

The identification of the schemas with the global mappings reveals the engagement of the neurocognitive information in every referential act, whether accompanied by external expressions or by only being thought. This is one more species of evidence for our neuronaturalized Fregean view that reference takes place only when there is information about the referent at the speaker's disposal, either already possessed or as incoming from the communicative situation.

I want to use the first identification above, that of concepts as schemas, to revive the old notion of *individual concept*. Individual concepts can be taken either in an ontological or in an epistemical sense. Taken in the ontological sense an individual concept contains all the correct information of the referent of the proper name. This sense is, if not the same as, at least very similar to Leibniz' complete concept of an entity that contains all true facts about the individual it is the concept of. In the epistemic sense individual concept consists of all the information one possesses about the referent of the proper name in question, but not all the information of the referent there is to be had. Of course the individual concepts in the epistemic sense do vary from person to person; trivially, in the case of humans, the individual concept that a particular individual has of herself is different from that of someone else's. And many people do not possess that concept for another trivial reason: they do not know anything about the individual and have never been in acquaintance with her. I am interested only in the epistemic individual concepts here.

Now if a schema, or a schema assemblage, as a global mapping contains information about a particular individual, who is the bearer of the proper name "NN", that schema contains all information that one has at his disposal in identifying that individual as the referent of that name. I propose that this schema *is* the individual concept one has of the individual NN in question. It follows from TNGS that the individual concepts are subject-bound and informationally varying, because they are the selectional neural processes of the global mappings of objects and persons. But this is no drawback, for it

was shown earlier that the informational variance of the senses is perfectly admissible feature of the Fregean theory, and one that he acknowledged.

So I take the Fregean senses of proper names to be the individual concepts of the bearers of those proper names. Hence I make a third identification, in addition to the two earlier (concepts as schemas and schemas as global mappings). This needs some qualification, though. The senses as the individual concepts, when they are activated as schemas, i.e. as global mappings of the individuals, are usually only partial as tokens, I think that this is the most typical mode of the senses of proper names to be entertained. For though we can take the sense to be the whole global mapping about an individual a speaker possesses, in distinct referential occasions different parts of the information about the referent is activated depending on the context, i.e. on the topic of an exchange and the information mediated. As I explained in the last chapter, proper names as pointers, now seen as entrances to the global mappings of individuals tend to activate larger portions of the global mappings than descriptions. But when descriptions are used, different descriptions work up different pieces of the information about the referent the speaker and the hearer possess; and these processes are not informationally uniformly the same in every occasion the referent is being talked about (if ever). Therefore differential informational "leakage" takes place that affects the content of the particular use of a name or a description. Some subschemas that were activated in the previous occasion may get activated again, but they may have changed informationally to some extent in respect to that earlier occasion. Moreover we should not forget that we tend to forget things. This is in itself always a change in the global mapping, hence in the individual concept.

This third identification of the senses of proper names with the (epistemic) individual concepts of the bearers of proper names is not entirely new. There are clear antecedents of it in for example Rudolf Carnap's and Alonzo Church's semantic theories. Also Linsky writes about particular individual concepts as being equivalent to the senses a speaker has about the referents of expressions (Linsky 1977: 72). My point is that by the help of the neurocognitive framework and the specific constructs it provides we can *unify* extensionally many distinct notions. That unification yields also a remarkable dividend in the form of modernized scientific explanation of the Fregean senses of proper names. I consider that to be a progressive step, and this for two reasons (at the very least). 1) Frege's semantic view, being an important part of his legacy, is defended in the way that opens up new vistas for research in the current semantics and theory of reference; 2) unifications are good in science, they provide unexpected explanatory gains and wide range of implications, and even lead to new research subjects. In this particular case my humble hope is that the way I have connected the neural and the cognitive levels to each other will prove its worth in further studies. For taking into account the large-scale nature of the schemas, and the schema assemblages, as the global mappings, their relevance is not restricted only to the issues in the theory of reference. For instance it obviously leads one to ask the question of the nature of the concepts in general, a field that has recently become the focus of keen interest in the philosophy of mind and in cognitive science. (Especially the involvement of the emotional and the motor elements in the global mappings points immediately to the enlargement of our current concept of concepts.)

If my identification of the individual concepts with schemas, hence with global mappings, is correct it follows naturally that the individual concepts are likely to

become related to, or sometimes even parts of, other schemas and individual concepts. As remarked above schemas can get embedded within other schemas by recursive operations, or part-whole forming processes, between the global mappings. This results in the schema assemblies which Arbib and Hesse postulate. These processes can be of short-term nature, restricted to changes in the synaptic activity states of the neural networks involved. But it seems that in many cases they result in a long-term assemblies due to the gene expression-related morphological changes in the synapses in the cortical networks and in hippocampus, i.e. to the pieces of long-term memories being parts of the individual concepts. As an example take again my individual concept of Josif Stalin (partly presented in chapter 6). Many of the subschemas of my schema of Stalin form assemblies with other (sub)schemas of persons, events and concepts about, and related, to the historical period of Stalin's life. This is a general feature of many individual concepts, so I try to make it evident in the following somewhat more enriched presentation of my individual concept of Stalin. The presentation is also affected by the contextuality effect that I argued is part of the senses of proper names. In this case it manifests itself in the way that the following presentation of my individual concept of Stalin contains some of the main features and events of Stalin's life. If the topic were, say, the Great Patriotic War or Stalin's relationship with Lenin, quite different features and events would be presented. This contextuality effect was already explained in the last chapter by the differential focusing of the portions of the senses. (The "... " marks the connections to the information about the mentioned individuals.)

STALIN

"Stalin" since 1913 [the main proper name pointer]; "Josif Vissarionovich Dzugashvili" the birth name; "Soso" as a child, "Koba", after a Georgian hero in a story [other proper name pointers]

-born officially 21st December 1879, really 8th December 1878, at Gori, Georgia, Russia

-went to a seminar to become a priest, but was expelled because of his political activities

-met Lenin first at Tampere, Finland, December 1905 at a Bolshevik party meeting

-V. Lenin:...

-Tampere:...

-Finland:...

-Commissar of Nationalities in 1917

-second wife Nadezda Alliluev (-first wife Jekaterina Svanidze

-had a son named Jakov

-killed himself in a German POW camp 1943)

-had a son named Vasili 1921 and a daughter named Svetlana 1926; killed herself at morning 9.11.1932

-*vozhd* of the Communist Party and Soviet Union since 1929 after ousting Trotsky, Kamenev, Zinoviev and Bukharin; General Secretary of the Communist Party of SU

-Communist Party of the SU:...

-SU:...

-L. Trotsky (Bronstein):...

-L. Kamenev (Rosenfeld):...

- G. Zinoviev (Radomylski):...
- N. Bukharin:...
- started the first *pjatiletka* in 1929
- let few millions of Ukrainians die of hunger
- worked closely with Molotov
 - Molotov (Skrjabin): shared a room with him in St. Petersburg 1912; worked in *Pravda*; became the foreign Minister of SU in spring 1939
- started the massive campaign to industrialize and dekulakize SU in the beginning of 1930s
 - hated peasants and farmers; first *pjatiletka* 1929-1933; famines especially in Ukraine with millions dead
- originated the purges of the higher party, economic and Red Army echelons as his targets in particular in 1936-1938 in the show trials; but also lower level party functionaries and other "enemies of the people"; signed at least hundreds of liquidation quotas
 - Kamenev, Zinoviev, Bukharin
 - marshal M. Tukhachevskii
 - NKVD
- G. Iagoda (Eiche):...
- N. Ezhov:....
- L. Beria:...
- negotiated with the Finnish envoys in October and November 1939; as the result started the Soviet-Finnish War ("The Winter War") 30.11.1939
 - J.K. Paasikivi:...
 - V. Tanner:...
- deliberately on purpose let Germany attack SU in 1941
- the supreme commander (at *Stavka*) in the "Great Patriotic War"; later assumed the title Generalissimus
- died 5th March 1953 in a dacha at Kuntsevo
- smoked pipe and tobacco; arranged drinking bouts in his dachas; vacations in south, mostly in Georgia and Abhasia; housekeeper Valentina Istomina (also his "unofficial wife" since the mid 1930s)
- vindictive and suspicious personality, perhaps even paranoid; occasionally cruel in personal relations; quick mind, exceptional memory (which he also valued in other people)

Perceptual information of Stalin (from photographs, news film clips, paintings)

Emotional information related to pictorial and film representations and knowledge about Stalin

In general, however, people seem to tend to possess somewhat similar individual concepts of the same referents. There are two main reasons for that. First is when the information content about the referent is "informationally poor", i.e. only a few things is known about him or her, but these are widely known among those who know anything at all about the referent, the similarities are perforce there. In the second case the concepts of the referent are informationally rich, i.e. so much is known about him or her, due to many criss-crossing communicative exchanges in a linguistic community

(and between communities, especially in our days of extensive attending to mass media), that the respective individual concepts tend to get informationally similar.¹⁴

I think that it is quite typical that most individual concepts locate somewhere between the poor and the rich ones, but maybe towards the poor end of the continuum. Especially this is so with the concepts the referents of which one is not acquainted, have not been *en rapport* with. But included are also those concepts the referents of which one has no special interests to know of. But that does not affect my point. These individual concepts are no less individual concepts than the rich ones.

Individual concepts identified with schemas, and schemas identified with the global mappings of the neurocognitive information about the referents, are, then, admirably suitable entities for the role of our modernization and neuronaturalization of the Fregean senses. Although I concentrate on proper names in this study, the unifying explanatory power of the schemas as concepts applies to Frege's notion of concepts and their senses (expressed by predicates). Second corollary is that the "pure" description theory of reference, according to which the Fregean senses are expressed wholly by linguistic means, by the definite descriptions, is too narrow theory. As we have been seeing, the individual concepts include such neurocognitive information which could not be linguistically articulatable, but which is functional nevertheless.

Binding by temporal synchronization

The last section concentrated on describing the kinematics of cognition underlying the reference of proper names, the neural structures and the connections supporting it. Now at last it is time for the dynamics and for the second seminal idea that I mentioned at earlier (section 7.3.1). Binding together of the sensory features by temporal synchronization of neural impulses to form coherent and unified perceptions has become the main subject of research, both in humans and in higher mammals (Treisman 1996). But

"...what next? What neural mechanism can be imagined to be sensitive to and recognize the presence of synchronization versus its absence? What mechanism can identify-i.e. *perceive*-the pattern of synchronized system activity as that of a particular external event? Whatever mechanism is proposed must function across a preconscious-conscious divide, still undefined in neural terms, for which no presently proposed paradigm presents a rational solution." (Mountcastle 1998: 376)

Essential as Mountcastle's questions are I do not think that the situation is so meager as he makes it seem. First of all the synchronization approach should be pushed further both theoretically and experimentally, for how else could we hope for progress?¹⁵ Most of the experimental studies of the neural synchrony have been performed using cats. For example, when a light bar is shown to a cat, the firing responses of the neurons in the visual cortex are found to be coherent and quite stable (even if transient). When the stimulus object is more realistic ecologically it is remarkable that in addition to the cortical (columnar) areas the synchronization has been found to take place also between distinct cortical regions, even up to the interhemispheric level. The mechanism of the synchronization can be considered as a generalization of the well-known Hebbian mechanism applied to the forming of all kinds of neural representations (perceptions and categorizations). According to the Hebbian

mechanism the presynaptic and postsynaptic sites of the neurons have to be simultaneously active to strengthen their mutual connections in the functional sense. Hebb's original proposal focused more on individual synapses, which is why I spoke of generalization of his proposal. Simply, synchronization means coincident and connected activity within and between neural groups in the milliseconds range (action potentials last typically 1 to 10 milliseconds). The oscillations in the neural activity patterns (like the by now notorious 40 Hz) that emerge from these processes may well reflect the fact that some of the activated representations are also relatively stable processes. Particular objects or topics "linger on", for at least seconds, even if only the different specific features of them become focused on. This may explain why our cognitive life is at least to some extent coherent in the extended temporal dimension and not just snapshot perceptions and thoughts (the coherence of which the milliseconds range synchrony takes care). When there is an overall topic of interest, as in an ongoing pointed conversation, thoughts stay focused. Or the objects in one's surroundings may play this stabilizing role when one is just looking. But when there is no such topic, as when one is just biding his time idly, different thoughts and images pop into one's consciousness. In the former case it might be that the synchronization within and between the relevant global mappings forms an informationally coherent and unified overall pattern, but in the latter case the pattern is more transient and weak in that it "disassembles" and "reassembles" continuously without any strong constraints. These processes go well with the recurrent architecture of the neural networks. That is because the processes are self-organizing and depend on the informational similarity and the functional proximity of the neural impulses. Or rather I want to suggest the stronger thesis according to which the synchronization goes well with the recurrent architecture because it *requires* that architecture. Even if input is processed in many distinct columns and areas, synchronization – hence representational-cum-informational stabilization - is achieved only through the vast arrays of the interrelated Hebbian connections in many levels between these columns and areas. If there were not these multiply reciprocal connections, temporary coincidental processes would be mere distributed and informationally unrelated happenings which could not offer us any neurally realistic mechanism for the explanation of coherent perception and cognition.

Edelman's approach requires also these temporal modes (Edelman 1989: 65). Therefore I want to suggest further (as I already pointed out in the beginning of the section 7.3.1) that the processing of the cognitive features of thinking is basically only more of the same as that which takes place in the perceptual bindings. Especially I want to suggest that the utilization of information in cognitive processes that contribute to speaking, remembering, problem solving and so on, are based on the detection and registration of the distributed synchronized events *by other synchronized events*. This overall process may be such in which "higher level" synchronical patterns emerge from the "lower level" ones and then take on and engage to guide the processing patterns of the latter levels.^{16,17}

As is well known there is no clean boundary between perception and cognition (and, for that matter, between sensation and perception, as well as emotion and cognition). Perceptions are constituted in part by those conceptual resources called to duty in every occasion when sensory material begins to come in. It is a platitude that what one knows and therefore expects affects what one sees ("seeing is believing"). This smooth merging of perceptual and cognitive factors may now receive the correct explanation

by the synchronous processing taking place in the re-entrant networks. But of course it is not *all* the same throughout: the distinct and differentiating features of the different types of the perceptual and the cognitive information are there during the processing and they bear their marks on the synchronous processes. (The reason may be due to the microcolumnar and columnar processing specificities.) Because the overall pattern tends to occur as holistically distributed processing, the relevant pieces of information are likely to be kept active during the particular perceptual/cognitive tasks. They so to speak find their proper places in the recurrent synchronous stream just by being in there at the right places at the right time. Most likely, however, the higher-level synchronous conceptual bindings are selective. By this I mean that not all and every one of the lower level synchronizations are engaged in or that they would become parts of the higher patterns of the synchronizations every time. What presumably takes place is that some of the lower level synchronizations are engaged (in varying extents) depending on the re-entrant connections and their synaptic strengths, which, remember, have informational impact (as far as we currently know). I am not claiming that there could not be other synchronizing processes simultaneously. It is quite likely that there are: one can think of something while looking at something else. That kind of diversity is only to be expected if the recurrent architecture of the global mappings is on the right explanatory track.

It may be the case that even if the synchronous processes within and between the networks through the different complexities and organizational levels *are* what we call mental representations, perceptions, concepts, emotions and plans, i.e. different kinds of schemas as global mappings, some of the higher synchronous patterns, even mappings of the global mappings, result from what - for want of a better term - I call "asynchronous side effects". That is, it may be that the dynamic collections of many of the lower level synchronous processes sometimes tend to generate asynchronous patterns among them. This may lead or directly cause a new synchronous overall pattern in the higher level connections. (Maybe something like this takes place when the so-called *Gestalt* switches take place, or when someone learns or experiences something that leads to profound changes in his schemas.)

Consequently the proposal on offer is that the synchronous activity explains why we perceive whole and coherent scenes, not patches of unrelated features or objects, even though the features are processed in distinct sites in the brain. And I further propose that the same mode of processing explains why the *cognitive*, or conceptual, representations, schemas as concepts, are efficient and likewise coherent (in their informational complexity). Maybe it should be noted that this mode of binding gets rid of any kind of homuncular tendencies because the processes of the recurrent temporal synchrony are postulated to *be* perceptions, imaginings and thinking. At first it may be difficult to comprehend that this is so because we are so used to see and think that when there is something to be seen or thought about, there is someone or something doing that. Maybe a comparison to the familiar and unproblematic identification from physical science would help to drive home the point. What is called light is electromagnetic radiation (sometimes waves, sometimes photons). That we humans are sensitive only to a narrow band of it does not make visible light something over and above the electromagnetic radiation (of that particular band of wavelengths). In the same way our perceptions and thinkings and other cognitive, as well as emotional functions, as coherent wholes are synchronical neural impulse processes in wide variety of areas in the brain, where the particular and task-relevant pieces of the

neurocognitive information reside. This holds whether they are only transient occurments or consolidated long-term items.

The neurocognitive framework for reference

Now we can begin to relate the recurrent architecture and the synchronical dynamics of the brain to our neurocognitive approach to reference. This proposal is a generalization from the studies and ideas presented above, and as such it is a somewhat speculative attempt. But that could not be helped at this stage of research if we are to take seriously the challenge of explaining reference from the neurocognitive perspective - which is the fundamental perspective regarding reference as a form of human action.

When someone perceives someone, or something, and eventually produces a linguistic expression referring to that individual, what takes place is in outlines the following course of events. Information about an object or a person gets processed in parallel and in an already *connected* manner because of the synchrony within and between the recurrent neural groups wherein that information resides. This is the individual concept as the global mapping (and schema). The neural activation spreads within the global mapping in a robust manner within and between the cortical sensory areas, starting from microcolumns where lateral inhibition by inhibitory interneurons directs it, as well as in the columns and eventually (i.e. well within the 500 milliseconds window) in the connected larger cortical sites, up to the relevant areas in the hemispheric lobes, most likely the (infero)temporal areas where information about objects and persons resides. This build-up is executed in synchronous manner by informational impulse correlations. In this same manner the situationally relevant information is bound together with the related activated memory representations of the object or the person (exhibiting the informational context dependence). The result of this global process is the particular schema or schema assemblage. As already stressed the amount and the activation of the "bits and pieces" of the memory information vary depending both on the situation and the internal constraints like one's expectations. As a further example of the latter factors the limbic regions partake by providing the emotional colouring and evaluation which can direct the selection of the particular linguistic expression (description, proper name, indexical or pronoun) which becomes as the result, the temporary end point, of this wholesale neural processing. This latter part of the process may extend to a few seconds, but with respect to proper names it may take much longer as in the tip-of-the-tongue phenomenon, when one just cannot recall the name although all the relevant information is in an activated state. (On this, see (Young 1998)).

If one is taking part to a conversation the activated networks, in which the information of the referent resides, keep that information "hot" and provide the background from which situation-relevant portions of that information i) get selected by extensive synchronous bindings and ii) used in one's subsequent linguistic remarks. This process can be likened to the figure-from-ground operations. This overall process is fluent, especially when it is established that the participants are talking about the same referent. I think that that reflects the large-scale engagement of the global mappings. Consequently I dare further to propose that this *is* what referring is in neural terms. Reference is not some cleansed seemingly out-of-cognitive-control-causal-relation from the referent to linguistic expressions. To use a slightly grandiose expression,

because reference is cognitive, it is also *primarily a form of memory*. To pre-empt an objection, it is easily seen that this holds also of the demonstrative references: the short-term and working memory systems are involved in these situations, in relation to the information the long-term stores feed into with. In the demonstrative referrals the lion's share of processing goes to the sensory and the higher perceptual networks, especially those that are activated by one's focus on the referent of the demonstrative. These are engaged synchronously with the activation of the basic schema of reference, this leading to the verbal expression of the demonstrative (and/or to pointing). But still these processes are executed by the short-term and working memory.

On the other hand reference being a form of memory may seem nothing but a platitude. But my view is different from some resembling views from cognitive science and philosophy of language. First, to blaze on it forth again, when put against the direct-causal theory, which does allow for cognitive representations associated with proper names but no mediating referent determining role for any cognitive constructs, I explain reference as being possible *only* because of the information the speakers possess and utilize. And as I have been arguing that mode is cognitive. The other difference is that I do not accept the sort of rather linsey-woolsey language-like cognitive constructs as lists, encyclopedias, files, dossiers and such, on which the mechanisms of search-through-and-hit would generate reference. Besides being too metaphorical this mode of operation is also much too stringent. Judged from what we currently know the only linguistic items involved are the end products of the neurocognitive processing: the neural representations of the linguistic items and the vocalizations (and other physical patterns like signs). Thirdly, even if some language-like representations are involved in the internal processes, the long-term memories consist mainly of cognitive, perceptual and emotional information, so the models of "pure" language-like cognition would still remain too restricted.

Reference as temporal phenomenon and as species of memory can also be supported by a closer look at the very causal links that the direct-causal theorists stick to. Picture to yourself what really makes the work, what retains reference, in the causal chains arising from the links and nodes. The nodes are occasions of communications between people, consulting sources (books, archives, pictures, paints, photographs, whatever). The links are the spatiotemporally enduring information stores of the people (=memory information) and the sources. When the information that resides in the sources is not "tapped" it is not used. When it is tapped, like in conversation, referring takes place because of the temporal activation of the relevant information guided by the topic and context. In fact this is what creates reference in the conversation also; the exchange has also additional functions like checking that the participants are talking about the same referent. This mechanism of tapping the information is in outlines just what I meant earlier when I said that the referential links are made and retained by us. This holds good also for the other sources, for though the information is there, they are just dead letters or pictures if no one is attending to them, if no one is bringing one's informational resources to bear upon them.

We should try to get rid of the habit of taking reference as a static phenomenon. When the referential use of an expression has an existing object as its target, reference takes place if the senses (the individual concepts) single it out. That is the general picture of reference. But we may be far better off explanationwise if we can avoid succumbing into thinking that it is the target that has the essential role. Information is *of* the

referent, and *about* it; it derives from it, when it is correct (and even when it is not all correct). But it is the *utilization* of that information that enables us to refer. Referring is a fluent phenomenon and so much dependent on the varieties of the neurocognitive information that it is the *convergence* of it to a unique referent that needs explanation, for the modes of processing of the neurocognitive information about the same referent are not informationally invariant every time they occur. This feature is notable in the intersubjective level because speakers must aim at the same referent for the referential communication ultimately to succeed.

The essentiality of information to reference allows wider range and flexibility to reference research than the direct-causal theory could allow. I mentioned somewhat passing in section 4.2.5 that sense theoretically understood reference can cease and be re-established. This happens when all correct and sufficiently individuating information is lost, but afterwards some of that information, or parts of it or entirely new information, about the original referent is acquired and reacquired (maybe by being inferred from related sources). Quite the contrary the judgement the direct-causal theory delivers is that once all causal connections from the original referent are cut then references to that individual are irredeemably lost. But it seems to me that even without the sense theoretically justified verdict, if one is allowed a recourse to referential intuitions in the manner the direct-causal theorists use them, then this type of case supports the sense theoretic notion of reference over the direct-causal one. For it does seem that after the cut in the informational continuity, we will still be talking about the original referent when new information, or the old reacquired, is brought bear. So the direct-causal theory is beaten at its own argumentative game.

The intentions to co-refer that the direct-causal theorists have used in their arguments, albeit erroneously, are also explained by the global mappings and their interdynamics. There are two kinds of intentions relevant here, and being intentions, i.e. action-conductive, they both belong to the class of executive functions. The intention to communicate is a general intention. Into it we can nest the Gricean-like maxims, for example "being to the point", "being patient and friendly (to a certain extent, but no more)", "keep it simple" (i.e. do not complicate matters unnecessarily), and other such maxims that direct the fluent exchanges between people. This is what I called wide pragmatics (perhaps not an altogether happy term). Intentions to refer, including intentions to co-refer, are less general and more dependent on the particular communicative situations but that makes them no less directive of communication. However their successful execution requires typically more of the neurocognitive information. Looked at from the angle of the global mappings this nesting of the intentions is only natural. It is the responsibility of the (pre)frontal areas to direct communication, through all the levels in it, from the "little plans" upward (see section 7.4). As the global mappings are engaged flexibly by the recurrent synchronized processing, the interplay between the executive schemas and the informational schemas takes place with no gaps. Especially I would like to propose here that the basic schema of reference is one such large-scale combined schema, activated in every referential act.

I must say a few more things about the role of language in my neurocognitive framework. I think it is evident, judged from the cognitive level phenomena, that neurally represented linguistic items boost the planning and execution of the schemas enormously in that those linguistic items become "basins of attractor", to speak in

terms of the chaos theory (Freeman 1999). In other words they tie large chunks of the perceptual and the conceptual as well as the motor and the emotional information into schema assemblies, hence relating to one another the multiple schemas of objects, events and actions. In particular I propose that the referential relations *presented within* those schemas emerge as higher-order schemas from the specific term-object associations. Indeed, the very notion of reference itself is such an abstraction. Due to the hierarchical higher-level recurrent synchronical bindings, which form the global mappings (and presumably mappings of such mappings), which is the mechanism of the formation of schemas-within-schemas, the brains are able to represent representation relations. An especially prominent function that the recurrency allows is what Edelman calls "recursive synthesis". It denotes the mechanism by which the higher-order constructions (maps, schemas) are delivered back to the lower-order constructions. As far as I can see this recurrent "recycling" can lead to the formation of new functions altogether, in particular through the temporal constraints in the networks. The recurrent architecture combined with the synchrony is able to perform these mechanisms:

"These effects include its correlative, constructive, and associative properties: reentrant signaling can modulate the timing of neural responses across neuronal groups, generate new response properties, and influence plastic changes by making them sensitive to signals from many brain regions." (Tononi & Edelman 1997: 699)

Cognitively speaking this results in schemas being about other schemas. So also the executive and the motor schemas could become represented by other schemas, referential uses of linguistic expressions being the specific cases of these new functions through the basic schema of reference. And even doubly so: 1) the linguistic expressions are about something external and represented as external *within* the relevant schemas; and 2) the act of referring by a linguistic expression to an object *as represented* also mesh together in the schema construction, and become part of the robust global mappings involved in the reference of proper names. This could also contribute to the referential intuition with respect to proper names the direct-causal theorists have banked on. That is, reference relation as represented within the particular schemas, grounded on the basic schema of reference, contributes the intuitive feeling of the referential force of proper names as such. But as I explained earlier this is a sort of cognitive illusion.

Presumably the higher-order referential schemas are essentially the interactive parallel processes manifest in the frontal executive, the parietal spatial areas, and the relational representational functions. There are also indications that the prefrontal areas are more involved in the relational processings to the extent that they require schemas (Robin & Holyoak 1995). (Recall also the role of Broca's area in relation ordering (section 4.2.5).) Lesions provide evidence for the involvement of the prefrontal areas, especially dorsolateral prefrontal cortex, in the construction of relational schemas: lesions in the mentioned area typically result in impairments of the internal model construction of events and relations (Knight & Grabowecky 1995: 1358). The involvement of these frontal areas indicates strongly that they are responsible for the organizing and ordering of the schemas. In regard of what the delayed-matching-to-samples experiments have revealed about the role of certain frontal areas in the short-term (or working) memory tasks it may not be far-fetched to propose that those frontal networks keep some portions of the processed information in hold and then let them

proceed further in their appropriate order in the ongoing processing stream. (In this respect the function is similar to that performed by hippocampus in that the latter orders and forwards information to the long-term memory stores in the cortex. Maybe there are intimate functional connections between the two regions accounting for their functional similarity.)

This then is my neurocognitive framework, or preliminary theory, of the reference, in particular of proper names. (Of course it has implications for reference of other expressions also, most naturally the natural, artifact and technological kind terms and concept words in general.) How much more detailed it can be made is an open empirical question. Of course I will welcome every result in that direction, but it should be noted that the nature of the subject is such that it could turn out that not too much to that direction should presumably be expected. By this I mean that the very dynamics and the accompanying structural changes in the neural networks may be such that the specific forms of information that are embodied in them could not be specified even in synaptical terms (but more likely not in biomolecular terms related to the particular structures in the networks, like the influence of the intracellular second messenger cascades to Ca^{2+} concentrations and to resulting amount of gene expression to adjust the synaptic receiving areas). It could also be that the specifications at those levels are not needed; it may be that the changes that are required for changes in the senses (in the individual concepts) are accumulative: it takes many small impacts before the "structural-cum-functional dam" breaks resulting in new configuration and pattern of dynamics. If it turns out that way, so be it. The more general characterizations, i.e. the columns, macrocolumns, recurrent connections between various networks in the various areas and the more large-scale connections between the regions processing distinct kinds of neurocognitive information are already sufficiently detailed to enable us to begin to build an adequate theory of reference from the neuroscientific perspective.

Notes

1. I presuppose this metaphysical view, period. I will not explicate it, but suffice it to state here that it precludes the supervenience theses currently in vogue. But as well it eschews the type identity theory of mental states. If pressed I would state that it does not postulate any *types* of mental states as real existents, if mental states are individuated by their content - as is the current view. Mental state types are only instrumental, or fictional, or some such convenient coarse-grained taxonomical categorizations, even if not explanatorily wholly inadequate. (I am a "neural realist" with eliminative urge, hence "mental instrumentalist".)
2. I have been using the notion of information throughout this work. As I already stated I claim no originality with that at all, for the idea that it is information that turns the referential wheels (and identificatory thoughts) has been around for, at least the last 50 years among the *neo-Fregeans*. To mention but a few instances, its seeds can be found in Strawson's account (Strawson 1959); information-based thoughts is a core notion in Evans' framework (Evans 1982; see also ch. 11); Recanati's theory (of neo-Fregean direct reference) makes much out of information, and is a descendant from Evans' account (Recanati 1993); and Michael Luntley has argued for the view that the causal links being primarily informational explains their referential properties (Luntley 1999). But the

modification I made in the beginning of section 4.2.5 - there are no causal links *qua* causal, but only informational ones – seems to me to be somewhat of a contribution to this issue. (I do not mean that what are typically called causes but which are information transmissions is a new idea. It is not.)

3. I am not saying that only humans are capable of language. As I mentioned the basic schema of reference is a phylogenetic structure. Moreover conceptual information is not exclusively tied to the language abilities, as will be explained later in this chapter.
4. Hopefully I would be able to present detailed arguments for my proposed enlargement of the domain of semantics in some future work. My view seems to have close affinities with some phenomenological theorizing, mainly with Merleau-Ponty's work on the role of the body in the constitution of mental phenomena and contents. On the other hand my view seems to go "deeper within" biologically but also to be more extensive externally because it takes the whole human organism with its surrounding context into account as contributing to the constitution of content (remember the contextuality of the Fregean senses argued for in chapter 4). It would also be interesting to give an account of what I like to call "technological kind terms" partly in terms of the neural motor and action pattern information and schemas. The contents of terms like "scalpel", "microscope", "magnetic resonance imaging", and such seem to involve a lot of the "practical information", so our enlarged semantics could be fruitful in studying the practical features of science, both the experimental basic research and applied science. In general I think that this would be a natural next step, for the overly linguistic explications of cognition, models of scientific research being part of it, have already begun to give room for the non-linguistic factors.
5. Emotional and affective factors of content must not be left at the level of general remarks. But detailed models of their mechanisms and effects on the cognitive and the perceptual processes must be developed. On this score I should warn the reader not to conflate my proposal with the quite well known associative network model of mood and memory by Bower (and its likes) (Bower 1981). His model is of the old "nodes in a semantic network" variety in which both concepts and emotions are labeled nodes, which become activated (and are recalled). As must have become evident I regard the conceptual and the emotional information being distributed through many areas in the brain, and involving, on both counts, unarticulatable features (i.e. not node-labeled). Models like that of Bower's are too coarse-grained, too folkish perhaps, to represent accurately the actual dynamics of how moods affect cognitive tasks.
6. Frege also held that language is not necessary for the grasping of thoughts; but we humans must use it for that purpose. In regard to the emotional states this view amounts to a fact: it is evident that many species of animals have similar emotional states as we humans do, but the expressions of these need no language. Rather they are "directly comprehended". (No commitment to naive empiricism intended; emotions, feelings and moods, as well as their expression, must be cognitively learned to some extent.)
7. CREB is "cyclic AMP response element binding protein". For a review of the issue see (Yin & Tully 1996). I want to point out that even within single synapses the re-entrant/recurrent processing pattern can be found, albeit in a much simpler form (though the point is controversial (Mountcastle 1998: 150-2)). There has been much talk about marks of signals received by the postsynaptic neurons getting back to the presynaptic side.

A prominent role has been given to nitric oxide, NO (a gas!). It diffuses to the synaptic cleft after the transmitter activity and reaches the presynaptic terminal. What is the specific function (or functions) of NO is not yet known. It may be related to the informational characteristics involved in synaptic transmission or it may be just a kind of confirmation signal telling that the transmission is going on unproblematically (somewhat analogously to the confirmation signal one gets into one's mobile phone after sending text message that the message has become received). Anyway my point is to suggest that the recurrent processing pattern, as a unifying mode of processing, may be pervasive in the sense that it prevails even at the molecular level. About NO see for example (Larkman & Jack 1995). (A type of the glia cells, astrocytes, surrounding the presynaptic ends of the neurons can provide for another, or an additional, mechanism for the recurrent "micro signaling". They can uptake the transmitter released to the synaptic cleft and due to this may be able to detect its (relative) amount and signal to the presynaptic end to change the amount of transmitter either produced or released or both. But that is speculation on my part.)

8. Parts of the theory have also been tested by modeling them with supercomputer and robotic constructions (Tononi *et al.* 1992). This "synthetic neural modeling" approach differs from the connectionist modeling because of its much closer adherence to the real neural structures and principles of neural dynamics.
9. As told the processing properties of the circuits in the primary sensory areas are quite stable, but it is also well documented by now that if some networks lose the flow of their specific type of input due to a lesion in the peripheral nerves, plasticity occurs and other neurons sprout to the "dried up" areas.
10. The informationally linked schema activation seems to be responsible also for the typical effect of facilitation by semantic priming, a phenomenon that can take quite astonishing manifestations on some occasions. For instance when one encounters a conversation about *cosa nostra*, and then some time later, even days, is queried about chess openings, the very likely first candidate that comes to mind could be Sicilian Defence.
11. Experiments on categorization have shown that various features of different kinds of information get encoded as memory representations. In addition to the specific items and their structures, what is done with them, how they are memorized and in what context(s) and mood(s), all influence their recall (Shanks 1997: 121). These findings support my general suggestion of the complexity of the schemas as concepts, and as constitutive of the content of thoughts and mental representations.
12. Though the issue is debatable: see Jerry Fodor's and Ernest Lepore's joint papers, for example (Fodor & Lepore 1996a, 1996b, 1999). Here it is not the place to counter their arguments in detail; suffice it to say that they seem both substantially and methodologically begging the question in that they presuppose stringent metaphysical approach to content and concepts, not naturalistic one. But I could not resist pointing out that the identification of concepts with schemas leads to no troubles with compositionality, the hobby horse of Fodor and Lepore. Schema construction to form schema assemblages is compositional as far as the activation of the subschemas goes. The neural mechanisms of the compositions are at least those mentioned in the main text, most notably that of recombination. However, it could not be avoided that the schemas change to varying extents during the compositional processes. This seems to lead naturally to the

view of content similarity. But it can be argued that this view sustains everything that successful communication and referential use of language requires, so the strict content preserving compositionality need not be respected. The interactive engagement of the schemas, as concepts, explains why the demands of the strict individuation are off the mark. We can talk about particular schemas, like restaurant schema, but this only means that the schema as one's memory item gets activated anew in every occasion when information about restaurants (or related activities) enters one's brain, or is activated by some internal prop like memory. Schemas are not frames or such cognitive constructs which have "slots" into which the specific information gets put. The information the schemas contain is part of the whole processing stream, which make up the schemas as token processes. The frame part can be interpreted as that information which has become to form relatively robust core of a particular schema. Compositionality may not survive always, either. It is perfectly possible that we have many schemas that have, so to speak, dissolved their subschemas of which they were originally composed, and has become to functions as units.

13. A simple operationalization to bring this out would be to ask people to rate individuals according to some scale as to their "positiveness" and "negativeness" with respect to one another.
14. Michael Luntley observes that there is a convention according to which ambiguous proper names take the most famous individual with that name to be their primary referent (Luntley 1999). But it seems to me that there is no such *convention*. The extent of information about the most famous person (when there clearly such is) in relation to temporal circumstances just is such that the individual becomes (has become) the most obvious referent of the ambiguous name. For instance we possess so much of information about Winston Churchill and about the Second World War that in most cases "Winston Churchill" refers to him (the prime minister of Britannia during the hardest years in the war), and not to a later MP in the parliament.
15. Of course other proposals should be advocated also; see for instance the open peer review to (Phillips & Singer 1997). With regard to synchrony some concise presentations include (Shastri & Ajjanagadde 1993), (Llinas & Paré 1996: 8-14), (Singer 1996), (Simmons 1996: 220-8), (Mountcastle 1998: ch. 12).
16. W. Phillips has pointed out that hierarchical structures could not be formed in the same level of synchronization in which their elements reside without destroying the synchronization (Phillips 1997: 38). This is correct, and that is why I postulate, or speculate, about the higher level synchronization processes. In another publication we read: "...information that is conveyed as a number of separate groupings at one level or stage of processing may be conveyed as a single coherent grouping at later levels or stages." (Phillips & Singer 1997: 679). This is very similar if not the same idea as mine. (It may also have much to do in explaining compositionality phenomena in thought.)
17. Oscillatory activity patterns anywhere in the brain may contribute to the facilitation of the cognitive processings. At least in me this happens usually when I am walking: my brain usually pops out various kinds of little (and even a little bigger) thoughts and ideas (compared to, say, when I am reading). (In fact this note is one such result.) This neurophysiological explanation, if it is generalizable, of peripateticism - thinking by walking - seems to be somewhat puzzling, though. Should it not be that walking increases

the utilization of glucose and oxygen in the muscles, robbing them for the brain to use? And in the brain the need of them is in the motor areas. But maybe effective synchrony requires *less* activity? That is, maybe thoughts and associations keep popping up just because inhibitory processes loosen their grip because of the lessened metabolic activity from the normal levels. Or it could be that the oscillatory synchrony creates a relative functional divide between the motor (and somatosensory) areas and other areas, thus enabling the more efficient use of them all individually.

8. THE EXPLANATORY POTENTIAL OF THE NEUROCOGNITIVE THEORY

The classical problems

In this chapter we will see how and to what extent my neurocognitive theory of reference - modernized Fregeanism - explains the most pressing classical problems in the philosophy of language and mind. Because the theory is not yet a tightly knitted one, much more work needs to be done. But, as said, let us see how far we can go to begin with.

The identity statements

This puzzle is about the informativeness of the identity statements "a is b" (put in the formal mode) in comparison to the trivialness of the identity statements "a is a" (or "b is b"), when "a" and "b" in fact refer to the same individual. As must surely have become obvious by now I think that it is obvious that Frege provided the essentials with which to solve this puzzle by his postulation of the senses of proper names. The postulation has been extended and naturalized during the course of this work by the notion of neurocognitive information embodied in the individual concepts as a particular kind of schemas, and their activation in the referential use of language, as neurally recurrent synchronously functioning global mappings, or parts thereof, together with the basic schema of reference (which, as we saw, also contributes the "metaintuition" that something is being talked about (named "NN" or described by "the F"). So what is left for us here is to propose some specific mechanisms by which the differences in the cognitive significances of the identity statements can be explained.

When one hears or receives by other means an expression "a", usually a proper name or a description, information is transmitted to and/or activated in one's brain in the way described in the previous chapter. This happens also in the case when one acquires "b" but when that information is not linked to the information which "a" activates; at least not in such parts which tap the referents neurally represented. But when one hears the identity statement "a is b", the neural groups as parts of the global mappings get connected to each other, embodying the respective individual concepts one has of the referents *a* and *b*. This can be seen as a minor *Gestalt* switch. It may take place via, what I like to propose and call a "catalyst network", i.e. a mapping between the two global mappings. Because the activation of the neural connection happens fastly in the occasion the identity statement is acquired, it could be that it is the synchronization between the mappings involved that causes the connection.¹ That the catalyst network may also be quite extensive in neural terms (for example by the count of the synapses involved) is shown by the fact that there typically remains the memory feature which can be characterized as "what were earlier kept or taken as distinct things have turned out to be identical". That is, if there were not the catalyst network(s), the information of the two mappings would just have coalesced with no mark of the earlier held informational distinction. (I take it that this feature of the catalyst network explanation is more adequate than those given by some neo-Fregeans that do not notice that being the result of the combining of the "dossiers".)

The specific re-entrant catalyst network connections include most likely those going via hippocampus. (Or the informational connection could be made through the "spare synapses" I proposed above.) This is also a little speculative on my part but I could not help offering it. From experiments we know that long-term potentiation takes place in hippocampus. That is, strong presynaptic impulses cause postsynaptic potentiation which endures even days *in vitro*. It is also quite certain that hippocampus receives already object- and event-bound information from many sensory and association areas. So it seems possible that the hippocampal networks are able to combine the input, the content of which is for instance the heard statement "a is b", to the informational memory input coming from the association areas in the temporal cortex which are also activated upon hearing the statement. The processing loops in hippocampus seem to be able to retain the catalyst information in an active state so that it waits until the memory information from the association areas, prompted by hearing "a is b", gets there. In this way the relatively distinct global mappings of the referents might get merged and the knowledge that *a* is *b* incorporated. So in hippocampus not just the occurrent moment-to-moment bound stimuli from the external objects and events get processed, but it would also make it possible to recycle the earlier acquired and consolidated memories.

The tetanic nature of this process could be explained by the amount of the information in the mappings, i.e. the potentiation may be a mass effect. In more specific terms the neural catalytic process may depend also on the frequencies of the action potentials and thus of the strength of the postsynaptic dendritic potentials and their efficiency in the propagation of the signals toward the somas of the neurons. This creates no traffic problems because the neurons partake in the encoding of many informationally similar features (the so-called population coding). These changes relate to the synchronous bindings in relatively short temporal scale. Long-term effects are mainly caused by the neurotransmitters and peptides as modulators (as was mentioned in the previous chapter). Later the informational connections between the two global mappings, the individual concepts of *a* and *b*, may grow richer in the cortical areas due to the gene expressions caused by the second messenger cascades. The metamemory feature I mentioned (i.e. what were taken distinct things are now seen to be identical) may also be explained by this hippocampal involvement. For we know that amnesics with lesions in hippocampus are typically unable to consolidate, maybe even to form, memories from the short-term encounters. In other words they do not have the normal feeling of familiarity that can be expressed as "this is something I have been through earlier". (Especially revealing in this respect are those hippocampus patients who have learned a motor task or skill but who do not recall at all that they have been through even quite demanding training sessions after the injury.)

For symmetry's sake we must ask what about the reverse phenomenon when there are in fact two distinct referents but a person has only one global mapping? In this case one's schema of an individual changes to some extent depending on the new information one acquires upon the discovery that there are two distinct individuals. As an example, Arnold has seen a very attractive woman around quite often lately. Most likely he has observed that she likes to wear rather darkish to emphasize her blond hair. Maybe he has also noticed that she does not talk much and seldom smiles. But then a friend of Arnold's tells him that this enchantment of his is in fact identical twins. What happens at that instant is that to the global mapping about the blonde is added information that there are two of them, and that he has presumably been seeing

both of them, separately, without realizing this (and other such "inferences" that dawn to him on the spot). Now he has an enriched global mapping with the information that it is about two women, twins in fact. But Arnold could not yet tell the difference between the twins. In particular he does not know their names (yet), so his global mapping does not extend to the temporal area in which proper names reside (area 38). As it may happen he gets later to know both twins better, so presumably also their names. By then the original global mapping has diverged to two (partially) distinct mappings, to two individual concepts (even if the twins happen to have the same name).

Here we see in action the neural specification of the mechanisms of the neo-Fregean metaphorical postulations like the "labels" of "dossiers" and "files". The labels are the neural representations of proper names, and they "call up" (parts) of the information about the referent, also the processing of (parts of) that information leading to the tokenings of that label, the proper name. (So *pace* Forbes, a neo-Fregeans himself (Forbes 1990), "labels" are not metaphorical any more in the neurocognitive framework; in fact "label" is very apt label for the neural representations of proper names.)

As a nice example of the relevance of emotional (and in this case even endocrinological) information to semantics note that not only the affective information from the limbic networks is an intimate ingredient of the global mapping right from the start, but so is hormonal information: Arnold's amazement resulting from being told that the blonde is in fact twins would not have been so warmly quivering were this not the case, i.e. were not the sensitization of the global mapping so strong. The emotional contribution in the now distinct global mappings might change for instance by being diminished in the other. Or it might have been the case that though the original global mapping was "compact" in the beginning, even then Arnold may have had a slight tendency to feel more drawn to the blonde on some occasions than in others. That is, his limbic system has been telling the difference between the twins, presumably because it is connected re-entrantly to visual information of the subtle kinesthetic patterns, facial and mimical features, and so on. That is, the limbic system has been sensitizing these parts of the mapping, maybe by differential hormonal (peptide) modulation in the neural groups involved. The result may be that in the subsequent divergence of the mapping the emotional features of the distinct portions of the mappings became readily different.

Disambiguation

I mentioned the importance of the disambiguation of the homonymous proper names when I argued against Salmon's argument from content variance and his caveat of the relativization of the use of a proper name. Let us see how the disambiguation goes in my framework. Recall that in this case the speaker or the hearer knows to whom a particular ambiguous name refers, or who are the respective referents of homonymous names (though this is quite an idealized assumption). The first case differs from the second only because we have to allow for the possibility that there is only one neural representation of, say, "John Smith" in the speaker's or the hearer's brain. In the second case there is more than one. Let us take the hearer first. The disambiguation starts from the activation of the neural representation of the name and proceeds to the global mapping related to the representation of the name in the temporal lobe, and at

the same time the activation spreads through the global mapping, or only to those portions of it informationally relevant in the situation, which encode the information about *that* John Smith. Or if the name is heard for the first time, a proper referent may be generated by the guidance of the context. In the speaker's case we may take this process as reversed, the result being an utterance involving the name "John Smith". If there are two, or more, neural representations of "John Smith", each tied to its own global mapping, distinct from the other mappings about the other John Smiths the speaker knows, the process of disambiguation is basically the same.²

Propositional attitude contexts

Propositional attitudes are described by folk psychological locutions like "believes that", "wants to", "imagines that" and such. This is the intensional characterization, the linguistic one. The intentional characterization concerns the properties of the states themselves. But here this distinction is of no importance for I assume the current Gricean consensus that the intensional is derivative from the intentional, that a person's psychological and cognitive states cause and thus explain the linguistic properties. The problems the propositional attitudes generate have been extensively discussed. Usually they are approached by pointing out that two central principles of extensional logic break down in these contexts. The existential generalization states that from the fact that a certain individual has some property (*a* is F) one can infer that there actually exists something that has that property, that something is F. The substitutivity of identicals *salva veritate* states that two (or more) co-referring terms can be substituted by one another without that changing the truth-value of the whole linguistic expressions of which they are parts. However, a child may believe that Santa Claus is fat, but it does not follow that Santa exists, contrary to the existential generalization. And there may still be people who believe that Morning Star is a planet but not believe that Evening Star is a planet, in spite of the fact that Morning Star is Evening Star, the planet Venus. Here the substitutivity of identicals (i.e. their names) *salva veritate* is violated, for it is false that those people believe that Evening Star is a planet.

There has accumulated an industry of proposals to explain why the two extensional principles break down. It should also be noted that the proposed reasons why especially the substitutivity fails are many and not nearly uniform. Obviously I could not take up here those specific proposals in extension; fortunately that is not needed either. That is because virtually all of them more or less do agree that the problem, thus the explanation, has to do with cognitive information at the disposal of the person suffering from the "defective" beliefs (or other propositional states).

It seems to me that the explanation of the failure of the substitutivity is in effect the same as the explanation of the different cognitive significances of the identity statements. In effect the situations are the same. The expression "a is b" has cognitive significance in contrast to the statement "a is a" because some portions of the neurocognitive information a person possesses is linked to "a" and some to "b", but not to both. The substitution of "a" by "b", and *vice versa* does not go through because the global mappings related to the neural representations of "a" and "b" remain distinct.

That is, as explained they have not become merged through the catalytic information that *a* is *b*. When the person acquires this information substitution will go through.

The failure of the existential generalization is a little different. Here we need not have two or more distinct global mappings, but only one in which resides all the information a person has about, say, Santa Claus. In contradistinction to the substitutivity case the person does not think that there are two different figures the one of whom slides down inside chimneys and the other who keeps on repeating annoying "ho, ho, ho" greetings, but only one who does both. Moreover the person keeps on *purportedly* referring to Santa because what is lacking in his neurocognitive repertoire is the information to the effect that Santa does not exist. Also as long as he believes in the existence of Santa the existential generalization does not fail *for him*, but it fails to hold nonetheless in the extensional sense because it is known that Santa does not exist.

These explanations are quite easy to come by but with the easiness comes a deeper insight. If the only piece that makes a difference is the knowledge that Santa does not exist, we immediately see that there is not any qualitative distinction between genuine and purported reference. The reason is that our subject keeps on purportedly referring to Santa as long as he does not become to know the sad news about the non-existence of Santa. Indeed we all could have used "Santa" that way, i.e. to have believed that Santa existed (but no one has ever seen him). This also supports my hypothesis that reference really takes place only in our brains, by being a form of memory in the form of neurocognitive representations. To give the most accurate neural account with respect to the existential generalization would require laying out the global mappings that encode and execute negation operations in the thought processes. But that explanation still eludes me, as it does everyone else. The only hunch I can give is to look at the frontal areas and, at the cognitive level, to the development of the concept of negation in children. But the informational part, the contents of our particular subject's Santa beliefs, is explained in my account by the emergence of the re-entrant synchronous patterns encoding such information.^{3,4}

Frege held that in the propositional attitude contexts the *Bedeutung* of that-clause is not the customary *Bedeutung* but the *Sinn*, the customary sense (the thought the clause expresses). In the attitude contexts the referring terms are not referring primarily to their customary referents. (As it is usually said the semantic innocence is lost.) This view has generated a controversy, for on the one hand Frege seems to be right: if one believes that Hesperus is a planet it does not follow that he also believes that Phosphorus is a planet, even though Hesperus is Phosphorus. (The neo-Russellians and Millians deny this. But I already argued that their reasons are i) *ad hoc*, ii) they have to invoke the Fregean factors, and iii) the singular, or Russellian, propositions are suspicious postulations when it becomes to their explanatory potential in relation to our cognitive practices.) The object of the ascription of a belief is one's thought, not the customary referent of the thought, or the referents of the sense elements in it. But on the other hand it seems as obvious that the ascription of the belief is about the customary referent also, the planet Venus in the example; so the semantic innocence seems not lost altogether.

I really have nothing new to say about this problem. (Or problems, for it seems that the basic problem is that we are not quite sure what precisely is the problem with the propositional attitude contexts – or whether there is *the* problem; see for example

(Sbisá 2003: 156.) But that is just as well because I think that Graeme Forbes has dealt with it adequately at least in outlines (Forbes 1987, 1990). So in the following I will just go through his general argumentation. Take the sentence

1) Ralph believes that Marilyn Monroe was murdered.

Forbes regiments this as

2) $B(\text{Ralph}, \langle \text{Marilyn Monroe was murdered} \rangle)$

(I have changed corner quotes to “<>”.) Forbes says that 2) is wholly specific about the content of the proposition Ralph believes (i.e. the Fregean thought). That is, we use the same expressions in the ascription as the ascriber does. But when that is not possible we loosen the degree of the specificity of the content we ascribe. That is achieved by the locution

3) Marilyn Monroe is someone whom Ralph believes to have been murdered.

This ascribes a belief to Ralph about an object but partly also specifies the content. 3) can be further analyzed by

4) $(\exists \alpha)(P(\alpha, \text{MM}) \ \& \ B(\text{Ralph}, \alpha \wedge \langle \text{was murdered} \rangle))$.

Here "B" is a belief relation between believer and a proposition, "Pxy" means "x is a mode of presentation of y" and "^" is for the manner of the combination of the senses which the "<" and ">" specify. "MM" is abbreviation in 4) and has its customary referent, Marilyn Monroe. (One can infer from it, with the premise "MM = Norma Jean Baker" that Norma Jean Baker is someone whom Ralph believes to have been murdered. And this without ascribing to Ralph a belief that NJB was murdered, for Ralph need not have a mode of presentation/sense for NJB.)

But what about the case when the ascriber have different customary sense than the ascriber for the that-clause? If

5) Lois believes that Superman can fly and Clark Kent can not

can we identify the Fregean thought, or the proposition, the that-clause expresses? It seems that we could not. For from the ascriber's perspective "Superman" and "Clark Kent" have the same sense (assuming that the Superman story is true and that the ascriber, among many others, knows that). It seems that the ascriber have said something true in saying 5), but it also seems that he has said something contradictory because the names he uses have for *him* the same senses, though they do not have the same sense for Lois. Read from the ascriber's perspective the ascriber is trying to get at Lois' senses of "Superman" and "Clark Kent". But how *could* he achieve that for he may not have any epistemic or cognitive relation to them at all? So the ascriber's perspective is not captured by the belief report.

Fortunately the previous analysis goes through also in this case of the intersubjective variation of the senses of proper names. The sentence

6) Lois believes that Clark Kent can not fly

is analyzed as

7) Clark Kent is such that for Lois' *so-labelled* way of thinking of him α , $B(\text{Lois}, \alpha \wedge \langle \text{can not fly} \rangle)$.

This is Forbes' *logophoric* analysis of the attitude reports. Here the ascriber does *not* refer to his or to Lois' senses of "Clark Kent", so the analysis avoids the above mentioned problems. Instead the ascriber says that Lois possesses information associated with the proper name "Clark Kent", that information expressing the (negative) property "can not fly". In short, a sense of the referent is labelled by a proper name (or by some other expression used to refer to that referent) used by the ascriber, though not necessarily by the ascribee (though usually they tend to be the same). This obviously requires that the referent is the customary referent of the names, and relativized to the ascriber and the ascribee. This accounts for the semantic innocence of the propositional attitude reports. Note that from 7) it can not be inferred that Lois believes that Superman can not fly. The reason is that Lois need not possess information labelled by "Superman", or she may have such information but it does not express the (negative) property "can not fly" (believed to be true by her of Superman).

As I said I accept this analysis of the propositional attitude reports. The main reason is that for all I can see it accords well with the practice of the attitude reports. On the one hand we could not (at least not yet for the lack of proper neurotechnological means) access the ascribee's senses directly. On the other hand when we report attitudes we try to be as true to (what we take to be) the referential content of the ascribee's attitudes as we can. It could not be avoided that our own senses so to say come the way because – trivially – we communicate only by them. But that does not normally matter because the senses, though intersubjectively varying, do determine the same referents. And when it happens that they do not, and we discover this, we will try to sort things out; and if that does not succeed, we just *refrain* from ascribing any referentially determinate attitude to the thinker (with hedges, of course). That just is the way it is with the attitude reports. We are usually content with the reports that have the referents right and characterized semantically as closely to the ascribee's concepts as we could. It is rather pointless in fact to try to capture the "proposition" the ascribee entertains, because it contains variety of informational connections to other senses in the ascribee's representational system (his "surrounding story"). Note that some of these informational connections arise at the very occasion the ascribee entertains the thought in question. This means that it could involve contextual elements (as represented elements, though not necessarily consciously). As many theorists have observed this only complicates the search for the "correct" theory of the propositional attitude reports and ascriptions. To my mind this search for the general theory is doomed to failure for the very reason that we *could* not attain the richness of the sense elements activated in every particular occasion of entertaining a thought. Consequently the only thing that needs modification from my perspective is that Forbes uses the notion of dossier, as we have seen many neo-Fregeans do. Because I am opting for a neuroscientific theory of the senses I have replaced this notion with the notion of neurocognitive information in the above characterizations. (The global mappings are left implicit here.) But nothing else of essence needs to be changed.

However it seems to me that the motivation of avoiding losing the semantic, i.e. referential, innocence that the original Fregean view is alleged to bring with it is in general an overreaction. It is true that Frege said that the customary sense becomes to be the *Bedeutung* of the that-clauses in the propositional attitude contexts. But from this it does not follow that that sense, i.e. the thought, would somehow cease functioning in its normal way. Its sense components (though distinct for the ascriber and the ascribee) still usually specify their respective referents for *both* the believer and the ascriber. From the fact that sometimes when someone believes that Phosphorus is a planet it could not be inferred that he also believes that Hesperus is a planet it does not follow that "Phosphorus" would not refer to Phosphorus (=Venus) in the believer's dialect (or idiolect, if the proper name is such that only he uses). Neither does it follow that "Hesperus" does not refer. (If indeed they both do refer.) If the ascribee's sense of "Hesperus" happens to be such that according to it "Hesperus" has no referent, then that just is so for him. But that does not affect semantic innocence because most people, or the experts anyway, know that "Hesperus" refers. The poor ascribee can be educated about his ontology. But that we focus on the senses, on the cognitive information, the speaker possesses does not stop the referential function of those very senses in that occasion of focusing if the ascribee's sense is not defective. Neither did Frege deny that, or imply that it would. (Again: the prevailing of semantic innocence can be argued on behalf of Frege on the basis of his Aristotle note where he maintains that as long as the referent remains the same the variance in the senses is not devastating. Remember that the whole climate in *Über Sinn und Bedeutung* is related to the propositional attitudes. So I conclude that we can have our semantically innocent cake and eat the variance of the senses as well.)

Empty names

Related to what has been said above, the problem of the empty names is best dealt with now. The problem is to explain why the use of empty names, names that have no actual referents, is cognitively meaningful. In other words i) why have those thoughts cognitive content that underlie the uses of the names that do not have referents, and ii) why they are used as genuinely referring expressions before their emptiness is discovered? Take the sentence "Santa Claus can't come until we're all asleep." (Recanati 2000: 215) Santa does not exist but it would be extremely implausible to explain why the children go to sleep, if that sentence did not express a contentful thought. The problem of the empty names is especially pressing to those direct-causal theorists, the majority of them in fact, who maintain that no thought (in the sense of singular proposition) is expressed if the proper name has no referent.

It seems to me that the most common strategy to explain the meaningfulness of the empty names is by invoking pretense attitudes. According to that approach the speaker of the above sentence knows that Santa does not exist, hence knows that when her children express Santa beliefs, they do not express any singular propositions. Therefore she can engage herself in pretense when she utters the sentence. That is, she knowingly uses a name that does not refer but *as if* it refers to Santa. (And her children use it as a referring term but without any pretense on their part - and she does not want to disillusion them, not yet anyway.)

As I mentioned in chapter 6 Evans put forth a pretense account of the empty names (Evans 1982: 353-68). I criticized that account there, but to refresh our memory I will

state the main parts of my argument here. Evans account is motivated by his background commitments. He endorses Russellian thoughts, and *de re* senses of which the former consist. With respect to the empty names this means that because there *could* not be a Russellian thought expressed by a sentence with an empty name in it, because the object-dependent *de re* sense is lacking, Evans sees as the only plausible way to adopt the pretense account. But that move is basically flawed. When we use an empty name in ordinary communication we do not engage ourselves in any pretense even as long as we do not know it is an empty term. Quite the contrary, we use it as genuinely referential expression (like the children above). For example think of how "aether" was used in the end of the 19th and in the beginning of the 20th centuries. Physicists were talking about the physical properties of aether, its effects on the light rays and shortenings of the measuring rods. They had quite elaborate theories of aether's role and effects in electromagnetic phenomena. In short ether was a real entity for them. So it is easy to see that the pretense account is clearly beyond any acceptable plausibility that even a justifiably *ad hoc* hypothesis can be granted. In the Santa case there is an obvious pretense, but that case is irrelevant because the mother knows that Santa does not exist. Instead I claim that the real problem with the empty names arises when we use them without knowing that they have no referents. Therefore any explanation by conscious pretense is out of place. (And non-conscious pretense – what could that be?)

The argument just presented supports again my earlier suggestion that there is no qualitative distinction between genuine and purported reference of proper names *qua* referential expressions. For only after we have discovered that the names we have used are in fact empty will we engage in pretense (should any needs for that arise after the discovery). That is, because of the force of the basic schema of reference - related to the information about the proper name emptily used – we have taken it as natural that there is the referent. But what we could not appropriately do is to adopt the pretense stance as methodologically *ex post facto* in our theorizing and then claim that it has been operative all the time in practice. This move fails because the empirical evidence refutes it outright, but also because it manifests the violation of the epistemic symmetry principle I argued for earlier. The reason for this is that the pretense accounts smuggle in post-discovery knowledge as to the non-existence of the referent of the empty name to the pre-discovery situations. That is strictly prohibited by the epistemic symmetry principle.

Recanati discusses uses of fictional and metafictional sentences as well as empty names (Recanati 2000). He takes up the pretense account, and for most of the time manages to keep the issues separate without confusing the innocent uses of the empty names to any pretense uses. However his characteristically perceptive and acute handling of the tangled issues seems to leave him with respect to empty names. Recanati argues that when it turns out that there is no one to whom a proper name "NN" was previously used to (purportedly) refer, then the believer can no longer be said to believe the singular proposition, say, that NN is F, because there is no such singular proposition. Still the believer believes a general proposition that there is an individual with such and such properties, who (or what) is also F (Recanati 2000: 225-6).

Now if we hold on to the believer's perspective, as I think we should (at least as closely epistemically as we can), what *we* no longer can say is irrelevant in a crucial

respect. For as long as the believer himself has no indications to believe otherwise, he continues to believe "singularly" of NN that NN is F, for nothing has changed in his informational economy with respect to "NN". In other words he uses "NN" innocently as a genuinely referring proper name, even when it is an empty one. We see that in those situations only the believer's singular belief explains his behavior, verbal or otherwise, but the "general" belief does not. So Recanati's postulation of "pseudo-singular belief" as something in between the singular and the general beliefs is not required. It should be noted that Recanati does acknowledge the option that we "describe the believer as taking herself to believe a singular proposition...while in fact there is no such proposition..." (Recanati 2000: 226) But he misaligns this immediately to the pretense account (it is *as if* there is such an individual NN) in which the ascriber of the belief fictively ascribes to the believer a singular belief concerning NN. But especially with the case of the empty names the *ascriber's* perspective should be avoided as far as possible when we are making sense of the *ascriber's* behavior, linguistic behavior included.

So I think that the pretense account will not do. Here we should also register Frege's own and straightforward account on the *unproblematical* nature of the empty names. It is manifest in his statement that we *presuppose* the existence of the *Bedeutungen* when we use proper names referentially, and that we can be, and sometimes are, mistaken when so presupposing (Frege 1997a: 162).

The explanation my neurocognitive theory gives of the empty names is the following. The name in question is linked both to the global mapping, as the individual schema/concept of the referent, and the basic schema of reference. But usually not only that for the global mapping is linked to other global mappings that provide for the epistemic "surrounding story" and that story does not involve any piece of information that there is no referent. (Or if it does it is then not accessed by the speaker.) Therefore the empty names are used as genuinely referring expressions due to the engagement of the basic schema of reference and the information the related global mappings contain. Thus the empty names are strongly epistemically supported by the surround story. The aether case exemplifies clearly this feature of the innocent uses of the empty names (or terms in this case). Only when it is discovered that the name has no referent, i.e. when that information is fed to the global mapping related to the empty name will its innocent uses cease. That happened when the physicists (gradually) became convinced that aether is not needed because Einstein's special theory of relativity has no need for that postulation and it explained the electromagnetic and other relevant phenomena as well as the aether theories (plus that from it can be derived new predictions, such as object's mass times velocity of light in *vacuo* equals its energy content).

Fictional characters form another topic that has been quite much discussed. But I fail to see why because they are explained by the cognitive mechanisms of the construction of the purported referents of the empty names. The only difference is of course that we know that the characters are fictional, hence the pretense attitude suits the fictional contexts. (And if one does not know that he is talking about a fictional entity, my account of the empty names explains that, not the pretense account.) If that were not so, we would regard novels as awkward constructions, especially those which blend fact and fiction. As an example of the naturalness of reading novels note that the main characters are introduced in piecemeal fashion in most cases, just like we acquire information about real persons. This explains also why there are not any pressing

problems with the "lacking information", i.e. with the truth-values of such descriptions that attribute properties and features to the characters about which the story is silent: neither do we have anything approximating a complete knowledge of the real referents. And we individually know about some historical figures more than about many currently living persons we refer to. (It seems to me that this "uninteresting" view of mine about the fictional characters is very similar to Daniel Dennett's (Dennett 1991: 79-80, 365-6).)

Scott Soames, following Salmon's ideas, tries to save the Millian approach in the case of empty names and fictive entities (Soames 2002: 92-3). The crux of his account is that we have a sufficient *de re* contact with both future and merely possible but never actualized individuals. This contact is possible when we are able, in principle, to observe a sperm cell and an ovum that merge to develop into an individual, or that could have merged so to develop. In effect the Soames-Salmon account postulates singular propositions consisting of singular propositions of those not-yet-existing and the possibly-existing individuals to save the Millian approach. But even if we ignore the fact that the Millian view by itself does not require any *de re* contacts in the specification of the semantic values (the referents) of the thoughts (singular propositions), what Soames states is incoherent on the very Millian grounds, for it requires (previously or currently) *existing* individuals as the semantic values. In both of Soames' examples there are no really existing individuals (not yet and not ever, respectively). In other words, the ontological statuses of the individuals Soames takes to exist are individuals only in the sense that we imagine them to be really (actually) existing ones. But when something is imagined to exist, it is *imagined* to exist, i.e. its individuation requires conceptual-cum-informational factors. And these pertain to all intents and purposes to the Fregean senses (albeit very weakly individuating in Soames' cases). It seems to me to be nothing more again than a mere *ad hoc* postulation to speak about singular propositions being about singular propositions consisting of the possible "zygotic" individuals to save the Millian view.

This - pardon me - ontological free-wheeling is carried to its extremes in Soames' handling of the names of the fictional characters. For they are postulated to be "a special kind of real, existing object; they, too, are abstract objects." (Soames 2002: 93) But fictional *is* fictional, not really existing; fictional characters are imagined and/or conceptual, nothing ontologically stronger need be assumed. The Fregean conceptual means are the most that are needed for the individuation of the fictional characters, but that implies no ontological status to them in our modern approach. This handling of the fictional characters by Soames and Salmon displays clearly the philosophically frustrating feature of the platonistic view they endorse: the reifications do not add any explanatory power whatsoever, they only double our ontology by pretending ontological seriousness. (An apt case of the pretense account, indeed.) And what is more, platonism should be argued for, not presupposed, if one could not help using it. The obvious reason is that it is an extreme ontological view. (Though as my remarks here indicate, I take it really to be not an ontological view at all, but merely an expression of mistaking conceptual matters for ontological ones.)

The direct-causal theorists of the neo-Russellian and Millian varieties have proposed to solve the problem of the empty names with "gappy propositions" also. (These seem in effect similar to Salmon-Soames singular propositions containing singular propositions.) A gappy proposition is an abstract structure that contains a "gap" in the

place of the referent in the singular proposition. As I am suspicious over the notion of singular proposition, the much more I am over the gappy proposition. To keep on putting it bluntly it seems to me only a desperate attempt, nothing more. If the referent place of the proposition does not tolerate any Fregean mode of presentation why would there being *nothing* in that place be more tolerable? (And please do not say that “there is not nothing, there is a gap”.) In other words I just could not take this notion seriously. Plainly the gap does not contribute to explain a whit the meaningfulness of the thoughts about actually non-existent entities.

Finally, Stuart Brock argues against the actualized descriptions view as providing a solution to the problem of the empty names (Brock 2004). If for example the sentence like “Santa Claus does not exist” is analyzed according to that view it says something like “the *actual* jolly fat man...”. If so there is the problem that, because actualized descriptions are rigid and apply to the actually existing entities just like proper names do, it follows that we are to infer that Santa Claus could not possibly have existed. But I reject Brock’s claim that the analysis is the correct one. The actualized descriptions view should be understood to involve the claim that once it is, or becomes, known that there is no referent for the proper name to be about, the actualized descriptions analysis reads instead “the...that was incorrectly believed to exist, does not actually exist”. That is, the negative existential is the metalevel report about the earlier referential beliefs with respect to the proper name in question. And a thought expressed by a sentence, say, “Santa Claus is white-bearded”, once it is known that Santa does not really exist, becomes indeed Fregean “mock thought”, i.e. a respectable thought but known to be about fictional character. And that does not block, either, the sound statement that Santa Claus might have existed. So let us turn to the existential analyses.

The existential statements

I propose that the existential statements are plain *comments*; that is their role in discussions. The positive existential statement confirms that, when a speaker uses the name “a”, he in fact refers to *a* and that it is correct to refer to *a*, i.e. that *a* exists. (When one uses a description the situation does not change in any relevant way, as is easy to see: one only zooms in on that element of the sense of the proper name that the description expresses.) In a way the function of the positive existential statements is social; in this it pertains to the division of (referential) linguistic labour. One may use a proper name tentatively and ascribe some properties to its purported referent based on the neurocognitive information residing in one’s relevant global mapping. But that information may not apply, or fit, to anyone or anything, although the name might refer to an existing individual. When this inadequacy of the purportedly individuating information is discovered someone corrects the speaker by stating the positive existential, and maybe offers an adequate description of the referent if it is clear that the speaker was after that individual but did not know anything (nontrivial) about the referent. Or maybe there is just a disagreement over whether someone named “NN” exists or not, and NN is reported by an outside party “in the know” to exist to put the matter straight. It must be realized that in these cases the negative existential statement is already used implicitly. The person who corrects the wrong individuation made, or presupposed, by the speaker could have said that “NN does not exist” precisely because there exists no one *thus individuated*, and it may not be clear that the speaker was intending a certain person but had no adequate information about the referent.

The existential statements focus on the presuppositions of the referential uses of proper names (and other referring expressions) but as Frege held when someone uses a proper name referentially he normally believes that the referent exists. When the purported referent does not exist, the negative existential statement is usually ventured, thus commenting on the speaker's incorrect beliefs as to the existence of the purported referent. In the same vein the positive existential statement is used as a comment when it is clear that one is hesitant about whether the purported referent exists or not.

***De re versus de dicto* reference**

De re-de dicto is another distinction about which the philosophers (and the linguists) have debated for ages. The issue is not just how to characterize it adequately, with possible subcategories like perceptual *de re* reference, and what that has to do with the transparent-opaque distinction (see for example (Recanati 1993, 2000)), but whether the very distinction exists at all. Because the material is huge and still keeps coming in I propose to distance myself from it by first siding with those skeptical of the viability of the distinction and then giving few reasons for my stance. But first let me make an explicatory remark. The distinction between *de re-de dicto* thoughts on the one hand and *de re-de dicto* references on the other does not matter here. It is obvious that when the issue is the senses of proper names we are only interested in the thoughts. And from this it follows, as should be clear from all that has been argued during this work, that the referential speech acts inherit their referential force from the thoughts and senses. This includes the demonstrative referings, for as argued in the previous chapter the context is represented in thought and is therefore effective in one's referential thoughts via the basic schema of reference together with the perceptual information present in the demonstrative situation. *De re* reference is indicated when it is said *of* an individual *a* that it is such-and-such. *De dicto* reference takes place when it is said of *whoever* (or whatever) that he, she or it is such-and-such. But this does not yet say much at all, for it leaves open the relations (semantic and/or epistemic-cum-cognitive) the speaker bears to the individual in both cases. In the same vein it only adds to the problem when one talks about the ascribed propositional attitudes of another person. The reason is that it may happen that in those cases what is *de re* to the ascriber may well be *de dicto* to the person the attitude is ascribed to. So we must turn our attention to the semantic and epistemic-cognitive relations instead.

Usually *de re* reference is said to take place when the genuine, the actual, referent of the used expression is intended as the target. When a speaker says "the oldest player in this tournament can still beat many of the younger ones", that is a *de re* mode of reference with "the old player" if the speaker has someone particular player in mind (presumably by having been in perceptual contact with him). If the speaker intends to say merely that whoever is the oldest player, he (or she) can still beat many of the younger players, the reference by "the oldest player" is *de dicto*. But note that the description seems to subsume also proper names: if one says that "Viktor Korchnoi can still beat many of the younger ones" he may be using the name as the name of the grandmaster he has close cognitive-perceptual contact, or he may be using it as a name of someone called "Viktor Korchnoi" of whom he only knows that he is a top chess player. The speaker is deferring here but his *speaker reference* has also more of the *de dicto* status than *de re* because he relates only the general descriptive expression to the proper name "Victor Korchnoi". I do not know if this feature of the *de re-de dicto* has

ever been noticed, but I claim that it is genuine one, and the reason is the informational content the senses of proper names and descriptions contain for the speakers. So we have here the first sign that *de re-de dicto* reference may be a continuum of cases and therefore has primarily to do with the variety and amount of the information related to the referential expressions.

I prefer to restrict *de re-de dicto* readings of sentences and utterances to the actual world. The subject gets more complicated if possible worlds or counterfactual talk is taken into account. For in that case *de re* readings cross-classify the actual world ones. In the *de re* reading both "Aristotle" and "the most famous pupil of Plato" can both refer to the actual Aristotle (when the description is actualized). But "the most famous pupil of Plato" can also be read *de dicto* in the possible world contexts. In that case one considers any possible world separately and refers to whoever in that world is the most famous pupil of Plato, which of course need not be the actual Aristotle. (But "Plato" might still be referring *de re* in that world to the actual Plato). That is the predicative reading of the description that I made quite much out of in chapter 5.

I do not think that there is any qualitative difference between the allegedly two different modes of reference. As I already hypothesized it is only the amount and variety of the neurocognitive information that matters. The cases that are taken to be *de dicto* references the amount of that information is typically rather scarce and more conceptual than in those cases where one makes *de re* references. Thus in the example above the speaker may not know anything more about the players than what his referential statement expresses (assuming of course that there is the sole oldest player in the tournament). The *de re* references are usually taken to involve some kind of causally closer contact between the referent and the speaker. This is likely to lead to a controversy over the relevant types of the causal relations in question. But this I take rather to suggest that the causal relations as such are not sufficient to guarantee successful *de re* references. In other words, it may be again the case that those favouring the causal account of *de re* references have mistaken informational relations for causal relations. For example the speaker may have been in a visual contact with all the players, or he may know whom the oldest player is by descriptions drawn from a reliable source who has had the required type of informational contact.

But this second alternative again shows the marks of the fragmentation of the *de re-de dicto* distinction. How much is there *de re* and how much *de dicto* involved when one purports to refer by invoking his memory information acquired via some source, and when one purports to refer "on the spot", i.e. by generating a correct description on the top of his head? Is the reference *de re* or *de dicto* when one has forgotten from whom or where he originally got the information about the referent, even when his statement is of *de dicto* type by its linguistic clothes (by being, for example, a general type of statement as above and with no close causal connections, let alone direct perceptual ones)? What about the cases when the speaker has not been in some kind of direct causal contact with the purported referent, but that purported referent the utterance specifies is in fact the oldest player and he is referred to by his proper name? We see that it is not the causal relation *per se* that counts but the neurocognitive information and the interplay of it with the other pieces of the contextual information that contribute to the status of the referential act.

The individual concepts as the global mappings embody may contain different amounts and kinds of perceptual information. When an individual concept contains no such information, when the speaker has not been in any “close” perceptual contact (have not been acquainted with the referent), the individual concept is of general or descriptive nature. But that does not mean that the individual concept in question is not about a definite individual, especially if the speaker *can* identify or recognize that individual, if asked to. That is, I do not see it as being a necessary feature of the recognitional reference that there has been perceptual contact earlier; for it is possible to recognize someone or something by using only the conception one has about the individual. For instance if I have been told only that the oldest player in the tournament still can beat many of the younger ones, I am likely to recognize him among the players by his appearance.

In light of these questions it seems to me that if the *de re-de dicto* distinction amounts to anything explanatorily, it only marks the two ends of a continuum, not any qualitative difference. The actual referential uses of language are in many cases so invested in both perceptual (and/or perceptually derived) and non-perceptual conceptual information from the surroundings and from the memory stores of the speaker, which mesh to varying extents in the global mappings, that to try to find a principled boundary between *de re* and *de dicto* reference looks futile. Non-perceptual conceptual information should be understood in the way that the amount of the perceptual information is low and/or not of much relevance contentwise in the particular occasion. This caveat is in order because there are no strict perceptual-conceptual divide either between perceptions and conceptions - as the dynamical architecture of the global mappings shows and as is well known from the research in cognitive psychology and in philosophy. When the connections from the referent (and to the referent by motor actions) have been many, or few but informationally intensive, the referential act is quite secure and towards the *de re* end. But this may also happen when one’s knowledge of the referent is extensive and sufficiently detailed although one has not had any close causal contact with the referent.

De re can change to *de dicto* (and of course *vice versa*). This happens when the perceptual information the speaker possesses diminishes and is lost eventually or forgotten all at once, leaving only non-perceptual conceptual information (expressible by descriptions). It is during this extended phase of losing the perceptual information that the direct-causal theory also begins to lose its plausibility as it mainly relies on the perceptual information acquired through the demonstrative referential situations. In other words, when only non-perceptual information is left about the referent, it is open to any new “surrounding story” to become relevant with the result that the conceptual information determines a new but epistemically as genuine referent as the earlier *de re* perceptual information. (Why this still might seem “intuitively” askew was explained in section 4.2.5.)

Why, then, has *de re-de dicto* reference distinction been focused rather keenly? I think that the main reason is the failure to keep separate the first person and the third person perspectives. I take it that what matters to reference is the first person perspective in the explanations of reference (and the ascriptions of the propositional attitudes). What counts primarily is how the speaker conceives of the entities he purports to refer to, not how others, especially we as theorists, see things. That is, though one must adopt the theoretician’s perspective on the issues under consideration, one must not adopt the

third-person stance exclusively in all issues as the only viable perspective. Here we have a clear case in point when the theoretician must adopt the first person perspective. And the first-person perspective becomes relevant only when it gets, so to speak, "first personalized", i.e. when the speakers involved get to know something from which it follows, them being aware of it, that they can change from the *de dicto* towards the *de re* mode along the continuum of the referential modes. It does not matter to the speaker whether or not he has been in some kind of close contact, *en rapport*, with the purported referent if he is not particularly interested in the nature of that contact. If he is and comes to know the nature of that contact (for example that he has been shaking hands with the referent unbeknownst to him that that individual is the referent of the expressions he is using), then from the third-person perspective, which the speaker can assume also, we can say that his reference becomes more *de re* than *de dicto* than it previously was. What is from our omniscient perspective clearly *de re* reference, may well be *de dicto* reference from the speaker's perspective for the lack of the more "involved" information, information acquired by some more direct means than for example by being told that the referent is such and such.

To put this in another way one easily falls prey to the violation of the epistemic symmetry principle in confusing the referential interpretation of the contents of the speaker's doxastic states with the referential interpretation of one's own states. It seems to me that many debates of the *de re-de dicto* references can be dissolved by resolving this confusion.

Of course it could be stipulated that *de re* reference always involves the purported referent, even if that referent is not the correct one judged by the content of the utterance in which the referring expression surfaces. Then *de dicto* reference becomes conceptual, or "general", in the spirit of Russell's theory of descriptions: *de dicto* reference is to who or whatever is such and such. (With respect to the descriptions invoked it may even be Donnellan's attributive use.) But I think that the few cases above of the interplay between close informational connections, especially perceptual, and the descriptive information *within* a particular context are sufficient to show that this stipulation is not explanatorily fruitful.⁵

Donnellan's distinction

Like the previous cases the distinction between the referential and the attributive uses of the definite descriptions has been discussed extensively since Keith Donnellan proposed it in the middle of the 1960s (Donnellan 1966). According to Donnellan a use of the description, say, "the murderer of Smith" in the sentence or utterance "Smith's murderer is insane" is attributive when it applies to whoever it is that murdered Smith. The referential use of that description takes place when it is uttered with a manifest intention by the speaker to refer to a contextually salient individual, as when one says that Smith's murderer is insane upon witnessing the disturbed behavior of the person who stands accused in the court. The motivation behind this distinction derives from the observation that in the referential use the description plays no part in the truth-conditions of what is said by the utterance (of which it is a part). The attributive truth-condition of "Smith's murderer is insane" includes that person who is Smith's murderer, whoever he or she is, and the utterance is true if that person is insane. The truth-condition of the referential use is independent of the description, for the description may be incorrect: the accused in the court may be falsely accused, but

the utterance is still true of him or her if he or she is insane (or even if he or she only behaves insanely there).⁶

I do not think that the distinction implies anything negative with respect to the Fregean sense theory of reference. This conclusion can be reached quite swiftly once we drop, as I argued we should, the widely shared but incorrect assumption that the Fregean senses are thoroughly descriptonal and wholly articulatable without any remainder in linguistic terms. If the senses are not thoroughly descriptonal, *even in the case of descriptions*, and even if it is accepted that Donnellan's distinction between the referential and the attributive uses of the descriptions provides sound counterexamples to the Russellian version of the description theory of reference (as seemed to be Donnellan's at least partial intent), the sense theory of reference remains unscathed.

But does not this argument seem a little bit too swift? Let us suppose that "the murderer of Smith" expresses all that someone knows about a certain person. Taken by itself this description does not determine any definite person as the sole murderer of Smith (assuming the murder was not teamwork like Caesar's death). The description can even be made informationally richer without it still determining the murderer. So there we have a possible case where the description expresses a sense, the rudimentary individual concept the speaker possesses, but it fails to determine the referent. In contrast the referential use of the original, not enriched, description picks up a definite individual. Is this not a clear counterexample to the Fregean theory of reference?

It is not. Notice, first, that the fact that the sense does not determine unique referent as the murderer could not be soundly used against the Fregean theory. The reason is that it misses the intent of that theory. The sense theory of reference by no means requires that every sense of every referential expression, however informationally meager it is, determines a unique referent. That would be unrealistic on the face of it. As long as the sense is not informationally sufficiently rich, i.e. when it does not even purport to single out unique individual, it is not in the proper domain of that theory as comes to proper names in Frege's extensive sense including proper names in our sense as well as the definite descriptions. As pointed out in section 2.1 the function and the concept expressions are a different matter. The individual concept may begin to develop around the rudimentary elements that can be given by general concept expressions (like "a famous physicist" in Kripke's well known example).

The second consideration that speaks against the referentially used descriptions being counterexamples to the Fregean theory is the following point. The referentially used description in the statement "Smith's murderer is insane" makes a sortal individuation to a human being. This expresses a more weighty part of the sense because it is understood that only humans can be murderers. (The fact that murderer in trial stand accused is obviously another background feature which contribute to the referential use of the description.) That this is implicit does not count against my analysis. The reason is the factor that creates the referential connection between this implicit sortal concept and the intended referent the speaker has in mind. This factor is the perceptual information that characterizes the intended referent, the one who stands accused in Donnellan's example. Earlier I argued that perceptual information can be part of the sense of a referring expression, part of the global mapping as the individual concept. In Donnellan's example it is just this kind of perceptual information that creates the connection from the intended referent to the description used. So it can be maintained

that the description "the murderer of Smith" literally understood does not carry any referential weight, or if it does that weight is due to the fact that the description becomes in effect a pointing expression, a demonstrative. When the person stands falsely accused, i.e. when the description fails to be true of him or her, the largely irrelevant referential role of the description is explained by the essential role the perceptual information assumes in the situation. In effect the description expresses only a lesser part of the sense that is about the accused and therefore its falsity does not matter. (It may help to understand this by relating it to what I said in section 2.6: false descriptions can be used to refer also in non-demonstrative situations.) In short, the referential use of the description reduces to all extents and purposes to demonstrative reference. (Including the "semi-demonstratives" like "that X" said in the absence of that X. But this case can be explained as anaphora – see the next section.) That is, if the description does not apply (or applies only to some minimally required degree, for instance that the murderer is a person, when the description implicitly involves a sortal concept), the pointing function it assumes has the force of the demonstrative "that": "*that* person (the accused murderer) is insane". So I am quite ready to accept this line of argument, for it does not have any negative consequences in relation to the Fregean sense theory.

It has been debated whether the referential-attributive distinction is semantic or (merely) pragmatic one. Here I only state my view. The usual type of argument for the pragmatic interpretation is that the referential uses of the descriptions are to be explained by the relevant aspects of the context, hence pragmatic factors. The attributive use is typically analyzed along the lines of Russell's theory of the definite descriptions and is semantic, having to do with the "standard uses" of the expressions (as Luntley, following Kripke, explicates them (Luntley 1999: 53). - Recanati has offered a new interesting solution which I will not consider here (Recanati 1989; 1993).) But if my approach is on the right track the referential uses are also semantic for they involve sense-constituting informational factors, even if being perceptual they do pertain to the context of use. To make this answer more efficient would require an argument to the effect that there is not any such strict distinction between semantics and pragmatics that those who invoke it in favour of the pragmatic interpretation seem to assume. (It seems that Recanati's approach provides the beginnings of such argument.) What I said about the explanatory non-existence, or at the very least relative nature, of the semantics-pragmatics distinction (in the sense of narrow pragmatics) in chapter 4 still stands, of course.

Anaphora

As is the case with the propositional attitudes and *de re-de dicto* reference, in reading the literature on anaphora it could not escape one's notice that there exists a respectable amount of explanatory hypotheses and counterexamples to all of them. For instance there are different kinds of rules proposed for scopes of quantifiers and orderings of expressions, with counterexamples flying criss-cross. I think that the main explanation of that quite chaotic state of the anaphora research is the very underlying cognitive mechanism of anaphora *itself*, a mechanism that the theoreticians have failed to notice, at least to my knowledge. Let me explain. That one writer can represent anaphoric sentences or utterances as examples and another writer other sentences or utterances as counterexamples to the first proposed analysis are *both* dependent on the fact that their examples come from real life situations, or from similar but imagined

circumstances, in which the sentences or utterances become sufficiently meaningful, and being to the point, only when the *context* of the sentences or utterances is taken into account.⁷ It seems to me that the criss-cross of the examples and the counters to them exemplifies the fact that language is used referentially much more creatively than that it would be expected to be used if we relied only on some narrow class of specific rules of anaphora. In other words, the examples in the literature are based on the already built flexible schemas of communicational situations, which permit the presentation of both real and possible situations in which the sentences or utterances are, or would be, semantically correct (even when they seem *prima facie* ungrammatical and/or semantically odd). The writers seem to concentrate so much on the sentences themselves that this automatic process escapes their notice.

A simple example of the hopelessness of the merely syntactical or grammatical analyses is given by the following sentences.⁸ "Eeva expects that she will get bed early tonight." Is anaphora from "she" to Eeva herself or to some other person? It will not help to try "The meeting was hell. Eeva expects that she will get bed early tonight.". The problem stays, for Eeva may be thinking of the woman who almost went off her head in the meeting. No answers will be received without more detailed information about the context. Another example, taken from the literature, is the sentence "Most men who own a car wash it on Sunday." (Evans 1985b: 138-9) Here "it" is supposed to be in an anaphora relation to "a car". But this is only one possibility. Another reading is that "it" relates to some other (kind of) object altogether, for example to a garage. (As could be the case with "Many new garages are built by the new neighbours. Most men who own a car wash it on Sunday." This sounds little odd but it is not anaphorically incomprehensible.)

It can be argued that my criticism misses the mark. For it is presupposed that the sentences or utterances under study give *all* the information relevant to the situation in which those sentences or utterances are used. Thus if one says that "Eeva expects that she will get bed early tonight", "she" could not but refer to Eeva. Likewise with Evans' example: "it" anaphorically refers to "a car". This is correct but only as far as it goes. For I can state my point anew: if the sentences give *all* the relevant information the negative feature of the typical examples is that they handle only very restricted and distorted situations. Most of the everyday conversations are much more extended, both spatially and temporally. An adequate handling of anaphora as a cognitive and *memory-based* phenomenon would have to take these cases into account as well, but those proposing syntactical rules do not ever attempt that. For example two persons see something luxurious enough to make them stare at it (and realize that they both do that). Then years later the other says "it was a beautiful sight" and the other one knows immediately to what the speaker refers, perhaps because the context of that utterance provides information which brings to the hearer's mind the memory of the original scene. Notice also that in this case anaphora is not even entirely a linguistic phenomenon in the sense that one expression would refer to another. For the two spectators may not have uttered a single word in the original circumstances. Here the anaphoric relation goes via memory representation to the referent. These types of cases may not be very common but that does not affect my argumentative point.

So I do not consider the typical way with anaphora to be a fruitful one to explain it. Understood more realistically the anaphoric uses of expressions are heavily context-dependent. One can read my argument also implying that there is no viable

explanatory distinction between sentence semantics and discourse semantics. In the same vein the distinction between intrasentential anaphora and intersentential anaphora is spurious, if given only a syntactic explanation. Instead I am proposing that, despite the immense variety of anaphoric relations discovered, the underlying and unifying explanation of them is given by the on-line construction of the schemas from the neurocognitive information the speakers and the hearers possess and share in the context. (On the complexity of the issues involved in anaphora and in its semantical, or information theoretical, treatment, see (Gawron & Peters 1990).)

Suppose someone says "Eeva is tired. The day had been hectic and she just wants to get much sleep." "She" refers anaphorically to "Eeva", we may agree. I want to propose that anaphora is adequately understood as a species of *internal demonstrative reference*, requiring working or short-term memory. Though, of course, the information it utilizes may be all drawn from the long-term memory as in the example above with the two spectators. First of all, anaphoric reference becomes possible only when one knows how to use personal pronouns but also other deictic and demonstrative expressions, so that my talking about internal demonstrative reference should not be taken to imply that only demonstratives are subsumed as the public expressions for the internally constructed anaphoric relations in one's working memory and, in many cases, in connection to one's perceptual processes. The label "demonstrative" is to mark the specific feature of the internal schema construction, i.e. that the outer linguistic expression of anaphora is the result of the internal schema representations of the anaphoric relations. We learn the mechanisms of these uses when we begin to learn language and the basic schema of reference develops (although personal pronouns are not the first expressions learned). When a child matures to the stage when he is competent with pronouns he usually has also his memory abilities sufficiently developed. In particular he is able to retain on-line representations of events active in the cortical-hippocampal recurrent loops, to which process can be added new relevant information from the contextual sources both from the external surroundings and the other internal global mappings. This on-line construction of the referential and the situational schemas takes the role of the context in which the demonstrative references are made. The internality of them can be clearly focused on when there is spatiotemporal gap between the original circumstances and the later anaphoric references. But in general anaphora is internal in the demonstrative cases through the working memory utilization of the schemas constructed "on the spot" about the on-going communicative situation.⁹

So what takes place in the speaker's brain when he utters the sentences about Eeva above is that his global mapping, the individual concept of Eeva, is activated (though not necessarily wholly). The perceptual representations of the events of that day during which Eeva got exhausted are linked to it. The pronoun "she" refers in that constructed internal representation to Eeva by being linked to the global mapping about Eeva and by being connected to the basic schema of reference which guides the anaphoric use of the pronouns, in relation to earlier expressions in the relevant situation. This way one can in principle account for all anaphoric referential relations, even the most complex ones. (To test this prediction one could envisage more and more complex sentences and see at what point one reaches the limit of one's processing abilities, i.e. one's comprehension ability of the anaphoric relations the temporal span of one's working memory allows.)

The involvement of the working memory in anaphora receives support from event-related potentials (ERP) studies. The so-called left-anterior negativity during 300 to 700 milliseconds from the stimulus onset is usually interpreted as representing the cognitive processing elicited when a word in a sentence refers back in the sentence (Kutas & Dale 1997: 223). According to my framework the activation of the global mappings (and presumably the inhibitory mechanisms involved) is responsible for this phenomenon. The potential changes recorded are mostly caused by the dendritic postsynaptic extracellular potentials of *synchronously activated* (pyramidal) neurons (Martin 1985: 644). But obviously there is no restriction to just one sentence because the duration of the short-term or working memory easily extends to several sentences and utterances. Presumably this is so because the schema construction is an embedding process, i.e. there are synchronizations within the larger mappings (and mappings of mappings); schemas are not constructed anew every time with the reception of new sentences, or with the production of them. During this neural anaphora construction the information that is linked to the referring expressions individuate the primary referent for the speaker. Also in the brain of the hearer the referring expression launches his neurocognitive schema construction and working memory retention (of even long since gone events related to the referent) by the now familiar recurrent synchronous mechanisms. ("I saw Eeva too. She looked really tired. I have never seen her in such a condition, except once. She was like a living dead.")

It seems to me that precisely because we are able to handle many sentences and utterances during a single occasion, and even larger chunks of discussions and contexts in the working memory, the syntactical and grammatical analyses are utterly on a wrong track (see also (Bach 1987: 224)). To put this suspicion of mine in general terms anaphora *is not* primarily a linguistic phenomenon at all. Sentences and utterances are just convenient but very condensed encodings of the informational processes that produce them. Of course there are structural similarities between them, but as far as the sentences and utterances are only encodings it does not make much explanatory sense to try to find only in them the grounds for anaphoric reference.

I want to continue a little on anaphora. I said that the spatiotemporal extension of it has not been attempted but that is not wholly correct, for there are steps towards that in an innovative and original work (Chastain 1975). Moreover it tallies well with my neurocognitive approach to anaphora as schema construction in the contextually guided working memory processes, in which the context can be both external and internal. Charles Chastain's seminal notion of anaphoric chains that are preserved during the evolution of discourse can be explained in my neurocognitive framework. This becomes more important because Chastain himself does not explain our ability to construct and retain anaphoric chains.

To give one example out of many in Chastain's work

"D59:#A squirrel came by a tree. A man fed it some peanuts. He had been eating them. The man was sitting underneath the tree.#" (Chastain 1975: 231)

("#" marks a boundary of discourse.) The anaphoric chains are "a squirrel – it", "a tree – the tree", "A man – He – The man", and "(some) peanuts – them". As we immediately see anaphoric chains are not necessarily (maybe even not normally, see later) restricted to single chains, but many of them can be handled cognitively during

the same on-going occasion of discourse. Of course such examples can be proliferated at will, but I want to point out that they can be extended spatiotemporally by "discourse jumps". These are possible because, as explained above, one usually does remember the original circumstances and the topics of these discourses. The participants - or at any rate those who know what the earlier discourses have been about and to what were referred then - can catch up with this discourse even years later and continue it as if nothing has happened in between. (Chastain mentions only in passing this involvement of memory (Chastain 1975: 251).) For example a participant to the discourse D59 above can later ask from the person who told the story what happened to that man who ate peanuts underneath the tree. If the person asked knows (or if it is just a story, if he can continue telling it) what happened to that man, he will most likely tell continuing the memory-based anaphorical chain.

An anaphorical chain typically starts by an encounter with the object(s) and person(s) in those occasions where the referential expressions for them are introduced. Namings are such typical occasions, but not the only ones, because one may be introduced to a new referring expression by being told what or who the referent is and what it is called. This involves new anaphorical chains. Hence the anaphorical chains can diverge as well as converge (as Chastain also points out (Chastain 1975: 267)). In fact the ways in which the anaphorical chains can be instantiated are many and varied, but they all can be explained by the Fregean sense theory of reference because they are informational. (Chastain also mentions explicitly this information-related feature of them more than once in his paper.) So anaphorical chains do not present any difficulties to my theory. Chastain does argue by anaphoric examples, however, that "denotationism" is false. Chastain takes denotationism as equivalent to the pure description theory of reference, not the sense theoretic account. Thus with respect to the latter theory nothing untoward follows from Chastain's arguments. The main reason for the failure of the denotationism instead is that it could not handle adequately perceptual contexts. According to Chastain what counts in them "is the causal pathway along which information passes from the object perceived to the perceptual context; it is this which determines the identity of the thing which is seen, heard, touched, smelled, etc." (Chastain 1975: 248-9; see also 254) What I argued about the contextuality of the Fregean senses in section 4.3.5 and Donnellan's distinction in the previous section enables us even to embrace Chastain's claim, although it should be pointed out again that the causal chain literally *is* the information that is transmitted.

We can see from Chastain's perspective that anaphorical chains form a large part of the referential uses of language. That is because many (most?) speaker references after the introduction of a name, or a borrowing of it, are anaphorical in the extended spatiotemporal dimension. This observation also tallies well with my proposal that reference is a form of memory. As described above, the schemas and the individual concepts the thinkers and speakers have constructed during the earlier referential communicative encounters retain the information that is neurocognitively necessary for us to be able to use linguistic expressions to refer to anything. Demonstrative, ostensive and indexical references are a little different, but not much. That is because demonstrative situations will occasionally come along in the already weaved anaphorical chain with respect to a particular individual. For instance if you have learned to whom the proper name "Sami Hyypiä" refers and have talked about Sami Hyypiä in his absence, but when you encounter him "in the flesh" then that perceptual-cum-demonstrative link in the chain, like when you say "there is Sami Hyypiä", owes

much of its referentiality to the very chain which have been retained by the information you possess about Sami Hyypiä.

Possible world semantics

The framework of possible world semantics (PWS) has been in vogue for many decades now. It is a general framework for many different approaches in studying intentional attitudes and reference (among other subjects). It uses tools from model theory and logics, mostly from intensional logics. But there are distinct views within it that try to remain close to the actual cognitive phenomena as much as possible. In this respect Jaakko Hintikka's approach is the most faithful one, cognitively speaking (Hintikka 1969, 1975, 1989). I could not go into all of the aspects of his theorizing, but two of them must be discussed here because they are the most germane to my framework. They are i) Hintikka's two methods of individuation/identification and ii) the neuroscientific explanation of them.

Hintikka has repeatedly argued against the explanatory viability of the notion of rigid designator. (See especially (Hintikka & Sandu 1995), though they ignore the causal component of the direct-causal theory.) His general point is that rigid designators do "yield" in the propositional attitude contexts. This means that the individual referred to must be individuated or identified in those contexts before we can talk about it as that particular individual. This is obviously contrary to Kripke's stipulatory view ("prefabricated individuals") according to which the individuals talked about are unproblematically the ones from our actual world. But according to the Hintikkian approach it is us who "draw the transworld heir lines" in the intensional and propositional attitude contexts, i.e. we individuate by *individuating functions* the "embodiments", or the "roles" the individuals play, in the courses of events under consideration in those contexts (Hintikka 1975a: 30). This view has a certain Kantian touch in it: in a sense the individuals we talk about, refer to, are our own constructions in that we individuate and identify them by our perceptuo-cognitive apparatus. This is very similar to my view from the essential role of the neurocognitive information in the use of proper names, and in the referential use of language in general. I would also like to point out that the problem of the disambiguation of proper names (equivalent to Salmonian relativization of use), which must be solved for the notion of rigid designation to be of any use in the case of homonymous names, is a close kin to Hintikka's arguments against the viability of rigidity as an independent phenomenon, i.e. as a non-cognitive feature of the referential use of proper names. The differences between our respective accounts are mostly found in the details of the mechanisms responsible for the individuation and identification of the referents.

It seems to me that the point above can be generalized quite obviously. I have been arguing that we need to presuppose some identificatory means, viz. Fregean senses, for the rigid designators to get off the ground in the first place. Now the two methods, or modes, of cross-identification provide such in general terms. The public method is obviously a part of the Fregean notion of sense, but so is the perspectival method. As was seen that is shown by Frege's Ateb-Afla example. Hence in effect Hintikka's two methods of cross-identification are not new to semantics and to the theory of reference when correctly understood, although in the contemporary climate in the analytical philosophy of language they certainly so seem to be. As Hintikka has pointed out in many occasions the perceptual perspectival method is similar to Russell's knowledge

by acquaintance. So I only point out that it is more appropriate to extend the historical span here to Frege with respect to both methods.

Let us take a closer look at the two methods of cross-identification and reidentification. He calls them public and perspectival or object-centered and subject-centered modes of identification (Hintikka & Sandu 1995: 274) (but earlier also physical or descriptive, and perceptual, demonstrative or contextual (Hintikka 1975)). Public identification takes place when a person knows or perceives *who a is*, i.e. *a* is identified by his known features and properties. For example E. Morse knows that it was Sarah Harrison who killed his own mother because the evidence leaves no other nearly as reasonable explanation. The public method of identification is akin to descriptive identification, but it seems to me that it can contain factors which could not be expressed by descriptions but which are well known publicly; for example, and most likely, aspects of the outlook of the referent. Perspectival identification takes place when one perceives *a*, i.e. *is* or has been in perceptual-informational contact with the referent. Thus Morse sees a rather slender brunette woman wearing a doctor's coat. This perspectival identification method requires putting the individual in some kind of coordinate grid or system within Morse's perceptual field. (It may be both ego- and object-centered.)¹⁰

Hintikka argues that the neuroanatomical discovery of the visual ventral and dorsal routes supports the existence of his identification methods (Hintikka 1989; Hintikka & Symons 2003). According to the common neuroscientific wisdom the ventral route, as it is called, goes via those temporal areas, mostly in the inferior temporal lobe, in which reside many kinds of information about particular objects and persons. That is why the ventral route is also called the "what system" because it tells what or who the object or person is one perceives. The dorsal route is involved in location processing and putting objects in spatial configurations, ego- and space-centered. In addition to the visual areas, the involved areas are found in the parietal lobe, most likely in the Brodmann's area 7 (and its subdivisions). It tells where and in what position the object is one is perceiving, hence the name "where system". In the same way in Hintikka's account the ventral system is supposed to take care of the public identification and the dorsal route of the perspectival identification (Hintikka 1989: 124-5).

This turning from the abstract intensional logic and model theory to the domain of neuroscience in search for the grounding of the explanations of the referential language use is very welcome indeed. But having said that and not to take anything away from Hintikka's original work on this matter it must be also realized that it is still more a venture than a developed stance by Hintikka. His account simplifies in that it keeps the two routes functionally too separate from each other. That is not in accordance with the current view in the functional neuroanatomy. As we saw above the different processes work in unison and need informational input from one another during their activity. That is the very idea of the recurrent synchrony in binding the features of the cognitive targets together. It is true that the global mappings, as individual concepts, contain networks from the temporal areas that partake in the ventral route. Moreover the contextuality of the senses of proper names has process-wise much to do with the parietal networks engaged in the dorsal route processes. But as I have been arguing the individuation and the identification of the referents requires also somatosensory, motor and emotional information (at least occasionally) outside the two routes for the full identification of the referent, be it thought about or

perceived. Therefore I fully endorse Maunsell's and Ferrera's point when they say "Although the what-where distinction laid out by Ungerleider and Mishkin has proved useful for describing the distinctions between the parietal and temporal pathways, it would be surprising if such simple terms completely described their functional differences." (Maunsell & Ferrera 1995: 458)

Hintikka tends to keep his discussion at the general descriptive level. For example we are largely left wondering how the "operationalizations of the meaning functions", which are at the core of his theory with the individuating functions, are supposed to be realized at the neural level. My point is that although it is fine to find the realizations (or implementations) in the ventral and dorsal systems, this has the effect of leaving Hintikka's PWS framework without much explanatory relevance. Another way to say this is that, despite the step to the right neural direction, the neuronaturalization of PWS still awaits us. (A very interesting subject by itself, though.) There are many steps downward from Hintikka's general PWS to the neurocognitive level proper in which the identifications take place. It seems to me that the most obvious way to close these gaps and to achieve the neuronaturalization is by assimilating (maybe even identifying at some stage) possible worlds/scenarios, as cognitive constructs, with the schemas and schemata of schemas (of objects, events and persons). This would have the consequence that we will reach the neural level quite quickly, through a unifying short cut so to say, if my identification of the schemas with the individual concepts and these with the global mappings is correct. Zooming in on the global mappings in question can then further specify the individuating and identificatory methods. With respect to a particular individual the global mapping, the individual concept of that individual, is engaged but normally only different portions of it are activated and utilized depending on the situation. Public individuation is typical in non-perceptual contexts, perspectival in perceptual contexts. The former type of individuation calls for duty more of the non-perceptual portions of the information in the global mapping from the temporal areas, the latter more of the perceptual information (in addition to what one acquires on the spot in seeing the referent) in the occipital and parietal areas. But when the emotional impact is not negligible the subcortical nuclei like amygdala and hypothalamus partake in the identificatory processing also. And in those identificatory situations that require motor information the frontal motor areas are engaged (think for example being in a dark room when you have to recognize some object only by the way it functions). As pointed out, and as is well known, it is quite unlikely that there are distinctively pure cases: perceptions are concept-laden, and due to the recurrent architecture conceptual thoughts about the referents involve perceptual information when the global mapping contains it. (Though Hintikka is aware of that, for he analyzes the sentence "*a* sees who *b* is" along the lines that there is a physically and publicly individuated *x* such that *a* sees that *b* is *x*, i.e. *b* is a "perspectival individual" (Hintikka 1975b: 18-19).)

To illuminate a little more how the neuronaturalization of PWS would go, the technical-metaphorical notion of diverging, or branching, and converging world lines in doxastic possible worlds can be explained by the separateness of two distinct global mappings (individual concepts) upon receiving the information that there are (at least) two distinct individuals, not only one as the person has believed. In converging there is merging of (at least) two global mappings through the neural catalyst network(s), which contains the information or has the effect that the person becomes to know (or believe) that what was believed, i.e. that there are distinct referents of the two proper

names, is in fact only one individual. In essence these processes (convergence and divergence of the world lines of individuals) were described in neural terms in section 8.1.1.

Hintikka argues that individuating functions form a subclass of the individual concepts (in Carnap's sense), from which it follows that the extension-intension (sense) dichotomy is not theoretically satisfying (Hintikka 1975c: 89-92). But individuating functions and individual concepts in *my* sense are functionally, thus explanatorily, equivalent notions: they are both *individuating*. The latter pick up the same individual in all relevant scenarios (epistemic and doxastic possible worlds) as well as do the former. That is, when the latter fail the former could not succeed, for both "give us the referents of our individual expressions" (Hintikka 1975c: 91). Hintikka's argument against the viability of the extension-intension dichotomy seems to be based on a misreading of Frege. Hintikka claims that there seems to be no inkling of the idea of functionality of the senses in Frege's writings (Hintikka 1975c: 89). But surely the Fregean senses of proper names are functional, i.e. pertaining to functions singling out individually particular *Bedeutungen*. One should not be misled by the partly metaphorical expression of the sense as the "mode of presentation" of the referent, or "ways of being given", as Hintikka evidently does when he distinguishes the Fregean senses from his individuating functions by claiming in effect that the former do not involve the individuating element (Hintikka 1969a: 105). It seems that Hintikka has even misconceived the function of the Fregean senses with respect to referring expressions when he says that "Substitutivity of identity is restored, in brief, not by requiring that our singular terms refer to the entities postulated by the so-called theory of meaning, but by requiring... that they really succeed in specifying uniquely the kind of ordinary individual..." (Hintikka 1969: 108). Clearly the singular terms do not refer to the senses in Frege's theory but to the very same ordinary individuals as they do in Hintikka's account. They do that *via* the senses, as they do that via the "individuating functions". So there is no relevant functional-cum-semantic difference between them whatsoever. Moreover the linguistic expressions for the Fregean senses are concept expressions "saturated" by arguments (in most cases the ordinary proper names) in the form of the definite descriptions (though, as we have been seeing, not exclusively). This linguistic garment enables us to see how the senses are functions in being individuating/identificatory. For example "the most famous pupil of Plato" is a function expression the extension of which is Aristotle (the *Bedeutung* being the concept *the most famous pupil of Plato*). ("_" is the most famous pupil of Plato", instead, is an "unsaturated" expression, i.e. functional expression the *Bedeutung* of which is a truth-value, true or false, when it is "saturated", i.e. in the place of "_" an expression of a proper name is inserted. And the sense of the sentence is the thought it expresses, i.e. that Aristotle is the most famous pupil of Plato.)

Hintikka also faults Frege's approach on intensions, i.e. the senses, on a par with faulting Carnap's approach on them. But it is very much arguable that Carnap's notion and grasp of intensions was not sufficiently Fregean, let alone the same as Frege's, for such equivalence being justified. In particular it does not seem to follow from Frege's semantics that the senses of proper names could not be restricted to epistemic and/or doxastic attitudes (possible worlds), unlike Carnap's intensions (Hintikka 1975c: 86-7). (I say "follows" because Frege did not discuss epistemic issues in terms of possible worlds or any equivalent notions.) Consequently PWS does not have an explanatory upper hand on the epistemic and doxastic attitudes. On the contrary it is only natural to

take the senses to be so restricted by Frege when he discusses propositional attitudes in *Über Sinn und Bedeutung*.

It may be that Hintikka is led to criticize Frege on the manner he does, i.e. to see individual concept as more general notion than individuating functions because he is interested in intensional logics and their semantics, whereas Frege was not. That is, if one puts possible worlds and their individuals first, it is easy to think that the Fregean senses and individual concepts are not sufficiently strict in individuating terms, for they seem to leave it open that in some worlds they do not pick up the very same individual as in some others.¹¹ But this is an unfair view for two reasons. First, and as I already pointed out, Frege was not interested in possible worlds. (Though he was interested in intensional concepts like propositional attitudes with respect to the senses.) Second, the Fregean senses can be taken to be sufficiently individuating with respect to the actual world. When it is realized that proper names are ordinarily used with the actualized-rigidified force, the senses *are* individuating also in the way that Hintikka's individuating functions work.

Notes

1. This seems to be tantamount to, if not the same as, the triangle pattern of re-entry (Edelman 1989: 65).
2. It can be a little more complex, as for example in the case when one hears "John Smith" but the disambiguation leads first to a wrong referent, which the context may soon reveal. Or this may happen because the neural representations of "John Smith" are so close to one another in the temporal area for proper names that they easily activate the global mappings as it were in a random fashion. The phenomenon of the divided reference, when a speaker refers to two (or more) individuals by one name at the same time because the information he possesses is not sufficiently detailed to separate the referents (unbeknownst to the speaker) is then a residual effect of the failure to disambiguate. Igal Kvat has discussed the issues (Kvat 1989). The cases Kvat presents as evidence for his account are mostly based on original encounters with the referents (perceptual and causally close ones). But the explanation of the divided reference stems from the informational "cluster" view given by Kvat, similar to the other neo-Fregean models in the literature, which allows him to enlarge his account to non-demonstrative cases. As with the file and dossier accounts, Kvat's account is at best another subtheory explained by my neurocognitive theory (for example Kvat's notion of cluster is, as is typical with such notions, left at the purely intuitive or phenomenological level). In particular it seems to me that his account of the divided reference can be given in neurocognitive terms along the lines in section 8.1.1.
3. By the way, it seems to me that the usual talk of extensional versus intensional (and intentional) contexts denotes not a qualitative distinction but a distinction of degree only. The intensional contexts are those involving beliefs and other propositional attitudes (but also alethic and other modalities). The extensional contexts can be characterized by the feature that in them the referents of the terms are at our disposal *qua* themselves. Hence "Morning Star is Venus" is an extensional statement, but when one believes that "Morning Star is Venus" then it is intensional because it is possible that one did not believe it, therefore the substitutivity of the two expressions could fail for one. However the

statement that Morning Star is Venus as such *is as intensional* as it is in the explicitly propositional context. The reason is that it is known (or at least "strongly believed") that the expression "Morning Star" refers to the same individual that "Venus" refers to. Here we see that the only difference is a difference in the epistemic (or doxastic) degree of the justifiedness or entrenchment of the identity statement. (Ditto for any other declarative sentences). That something is known makes it no less intensional, and intentional. This may have been lost in sight just because we take some statements as unproblematically true and widely known (as established facts). Therefore the propositional operator has been dropped. But it is there tacitly for it may happen that we will discover that even the most cherished parts of our knowledge turn out to be false. The tacit intensionality/intentionality is explained because knowledge is informational business, thus cognitive. And as all human behavior has cognitive element, no "pure" extensionality could ever be distilled from our knowledge. (The shortest argument here is the following: the extensional contexts are at least sometimes knowledge-based, knowledge requires reference, the latter is intentional, so is the former.)

4. In the discussions of the contents of the thoughts and the propositional attitude contexts Kripke's "puzzles" about belief ascriptions are often mentioned (Kripke 1979). Kripke's argument is that there is something wrong with the very practices of ascribing beliefs to other speakers in that they sometimes lead to contradictory results like that one says both that "X believes that *a* is F", and that "X also believes that *a* is not F". If this is so, Kripke states, the blame for these results can not be put on the Millian approach, and in favour of the Fregean one, as is usually done. But I fail to see what is supposed to be the force of Kripke's examples because they are rather underdescribed. Kripke fails to take into account that the belief ascriptions, as he gives them, can be naturally further specified with the result that the contradictory ascriptions do not follow. Especially Graeme Forbes has argued that this, in fact, is what our normal practices of belief ascriptions allow for, and even require when there arises the danger of contradictions in the ascriptions (Forbes 1990: 356-63).
5. *De re-de dicto* references have usually been discussed with respect to descriptions. But as indicated in the text it seems to me that there are both uses of proper names as well. *De re* use of a proper name takes place when the speaker knows who the referent is by a close connection ("acquaintance" seems to capture this mode of referential contact). *De dicto* reference takes place when the individual concept the speaker possesses about the referent includes only information that is not acquired by acquaintance, *or* when it does the speaker does not know that (as in the hand shaking example in the text). This last possibility seems to indicate that there is a continuum from *de dicto* to *de re* with respect to proper names also. But this should be no news by now because what matters to reference is the amount and kinds of information one possesses, not the causal contacts *qua* causal with the referent.
6. Note that the referential-attributive distinction seems not to be same as the *de re-de dicto* distinction. When someone thinks, upon seeing Smith's mutilated corpse, "Smith's murderer is insane" that thought is attributive in Donnellan's sense but also somewhat *de re* in the sense that there is a rather close evidential (causal/informational) connection from the corpse to the murderer.
7. To my experience it is rather common that the evidential bases of the theoreticians' claims are suspicious in that some sentences are judged ungrammatical according to the rules the

writer proposes, but the very sentences seem not to be ungrammatical, or misconstruals, at all; especially not if one can imagine a larger context which enables one to make perfect sense of the "ungrammatical" sentences.

8. Another example is the type of outright question begging in anaphora studies. As Jaroslav Peregrin points out, the formal semanticists are *not* really allowed to presuppose co-indexing to have taken place before they get to work on anaphoric cross-references (Peregrin 2000: 274).
9. For a short account of "world knowledge" (in my framework schemas constructed from the neurocognitive information, and other previously constructed and consolidated schemas) in constructing anaphoric references in mental models see (Garnham 1997: 160-1).
10. The characters are from Colin Dexter's completely superb, occasionally dejected and definitely irrevocable *The Remorseful Day* (1999).
11. That my reading is correct is supported by what Hintikka says in another paper: "Arbitrary functions of the former kind are essentially what many philosophers call individual concepts, while the latter, the narrower set of 'individuating functions' essentially represents the totality of the well-defined individuals we speak of." (Hintikka 1969b: 137). And as to the functional equivalence of the Fregean senses and the individuating functions, there is a revealingly innocent remark by Hintikka: "Admittedly some further questions may still persist, for instance questions pertaining to the nature of individuating functions naturally used in our ordinary discourse. I shall by-pass them here, although they certainly need further attention." (*ibid.*) They certainly do, as we have been seeing. Hintikka also speaks in quite many occasions of the *sorts* of criteria the individuating functions have to fulfill, like similarities between worlds (for according to his view the identities of the individuals may depend on the whole world) and the spatiotemporal continuity of the individuals. But these sorts of criteria are not banned from the Fregeans to invoke.

9. REFERENCE

On neuroscientific reference

Towards concluding this work a few general comments about reference are in order. Linguistic reference is the phenomenon that makes us able to talk about objects, persons, events, properties, relations and other things. Put in abstract terms reference is some kind of relation between the referent and the referring expression. But this is also a mere platitude that in itself tells us nothing that would be of explanatory value. Moreover it characterizes reference as static. The prevailing direct-causal picture of the reference of proper names (and the natural kind terms) is also such a static one. As we saw the relations from the referents to the uses of the expressions referring to them are highly idealized and tend to be taken as unproblematically and intactly given. The source of this staticness can be traced to the plain generalization from demonstrative references, and to the coarse-grained and idealized conception of the causal-communicative relations (as was argued in section 4.2.5). True, that the direct-causal theory takes into account the communicative events, reference borrowings gives it dynamical flavour but only to a minimally required degree: without that modicum of dynamic elements no theory of reference could be adequate for reference is a form of human action. So contrary to what the common view of the direct-causal theory seems to suggest, the picture which that theory draws of reference is not sufficiently dynamic. The main reason for this is that it does not take into account the essential roles that information plays in every occasion of referential language use. As I have been arguing it is only by studying these roles that we are able to construct *the* adequate theory of reference.

While accepting the abstract image of reference presented in the beginning of this section, according to which linguistic reference is some kind of relation between referents and expressions, it helps to conceive of reference as a continuum along the *distance* between the referent and the user of the referring expression. That continuum begins with the shortest distance, the null distance. This is the case when the user is the referent: uses of "I" to refer to oneself. From these cases the distance lengthens, when indexicals and demonstratives ("you", "this", "that", and such expressions) are used.¹ Another way to conceive of this end of the continuum is to realize that in the case of "I" its reference and context are the same, so that no referential mistakes are made. With the other indexicals and demonstratives there is always such a gap, therefore the contextual aspects must be observed. For example in the case of two very similar objects standing side by side, then individuating them must invoke their relations to other "landmarks" in their vicinity, or at least to some of their mutual relations (for example being left of the other).

Towards the other end of the continuum reside proper names and the definite descriptions (not in such Donnellanian referential uses which are essentially demonstrative like "that man drinking martini", or a proper name functioning as a demonstrative, as in "Eeva is sleepy" when one perceives Eeva). The role of the conceptual information in the sense of non-perceptual information is of paramount importance in these cases: references to the long since dead historical figures are representative cases. It is also typical of these cases that the context in which the referents are located becomes *internal*. Of course it is obviously understood that the expressions refer to the individuals in their contemporary circumstances, but what I

mean is that those circumstances, the contexts, are *represented* in the speaker's brain by the information and the schemas constructed of the information the speaker possess. As earlier described the global mappings enable us to represent the referential relations within other mappings (schemas-within-schemas). In many cases the represented circumstances *are parts* of the senses of the referents, even the individual concepts as global mappings). I argued for this nature of the Fregean senses in section 4.3.5 and in chapter 6. The cases toward this end of the referential continuum manifest most clearly the dependency of reference on information. By the same token these cases make the strongest case for the cognitive and epistemic nature of reference.

However it is not the case that in the two former types of reference the context could not also be internal; rather it *is* also internal, neurocognitively perforce so. Every perceptual situation gets represented in the brain. With "I" this is obvious, for so much of I-senses are internal; though of course the surroundings one perceives do occasionally accompany one's I-sense, i.e. one's brain locates and relocates "its I" all the time spatiotemporally. What is going on is that the context is represented in one's brain but at the same time it is also *represented as being the external context*. The best way I can formulate this is that *in* and *during* the process in which the context is represented it is also tacitly represented to the organism as being external to the organism. The secret of this may well reside in the dynamics of the mappings of the spatiotemporal global mappings (schemas-within-schemas). Also the basic schema of reference presumably plays a role in learning to distinguish oneself from one's surroundings. It seems, then, that especially here the theory of reference and the theory of consciousness could meet each other. (In addition to Edelman's account see (Damasio 1999) for a philosophically very intriguing account of the neural construction of representations within representations, or rather oneself being represented within a neural representation in which that being-within is also represented.)

One can maintain also that proper names are *minitheories*. When we use a proper name "NN" purportedly (and non-demonstratively) to refer we in effect put forth a hypothesis that the features and the properties encoded in the neurocognitive information, the sense as the individual concept about NN, will get verified in the communicative situations. When that does not happen we either drop the hypothesis ("NN does not exist") or modify it ("NN is such-and-such, but not so-and-so"). The verification of the senses of proper names is quite a common phenomenon. Not only the reference fixings but also the intermittent but continual contacts with the external reality keep the grid of the referential parts of language in check. But there is always another type of check going on. In the conversational situations in which the non-demonstrative (purported) references are being made, it is the "meeting" of the mutual information the participants bring to bear which checks the coherence and cohesiveness of that information, hence the sameness of the referents the participants are talking about.

This property of proper names is sometimes noted; for example Evans mentions proper names as hypotheses (Evans 1982: 392). I want to emphasize this because it helps us to realize more clearly that the use of proper names as a form of human action ties them to our overall being as the creatures we are as *informavores*. However there is a deeper cognitive-evolutionary point behind considering proper names as minitheories. For it is obvious that we, as language using animals, have evolved to the current stage

with its abundance of the scientific and other theories from the rudimentary information that contains accounts of the first referring expressions (or vocal grunts pointing other's attention to their mutual vicinity). In this way the referring expressions acquired rudimentary senses from the very start. That is how our perceptuo-cognitive apparatus functions; the later theories – including those of individuals via the senses - only build on that, with their rich structures, ramifications and all.

In the end of chapter 4 I maintained that reference is intentional in being cognitive and epistemic because it is informational. Now, after all we have been through in the last two chapters, I think it is sufficiently argued that *reference is primarily a biological phenomenon*. Reference takes place only within our brains. As I argued reference is a form of memory, when memory is explained in the neurocognitive terms. I do not think that the insistence on the general static picture of reference according to which reference is only some kind of general, even if “causal”, relation between the referring expression and the referent suffices for explanatory purposes; that relation must be dismantled to its functioning elements. With respect to proper names the relation is informational and intentional, and as we saw it is made up of many different neurocognitive elements, both in the transient sense as in demonstrative situations and in the spatiotemporally extended sense as the sources of the information concerning the referents from past. The basic mistake of the direct-causal theory of reference is then not so much that it insists on the general static picture but that it insists on a very simplified form of it: that there are causal links which safely keep proper names together with their referents. That picture we should reject. But I am willing to go all the way for this opens up a profound ontological consequence. Because reference is so complex phenomenon - both synchronically in that it taxes brain's neurocognitive resources and diachronically because the "nodes" of the information transmission routes are many and varied - maybe we should also give up conceiving reference primarily in relational terms. Or if one could not do that, maybe he should try and conceive reference as being a relation only in a secondary sense, the primary element in reference being informational. Of course reference as an intentional phenomenon has the connotation that it is relational (for after all in intending *someone* intends *something*), but it is right here that the modern Fregean neurocognitive approach helps to loosen the grip of the relational picture. The senses of proper names *qua* informational processes carry the main explanatory burden: they are the mechanisms of reference. The relationality of reference is only a residual effect, so to speak. As I speculated above, the abstracted and generalized representations of the referential situations – schemas-within-schemas – and one being conscious of that representing of representing might explain the relational grip (together with the demonstrative *Dasein*). Anyway I hope that I have succeeded in arguing for the attitude and the consequent approach that we not only can but should begin to see reference from the new neurocognitive perspective; a major *Gestalt* switch indeed.

What makes this a theoretical *Gestalt* switch is that the referential and the communicative phenomena do not change. The switch is instead back to the internal perspective, but a richer one which easily accounts for the features responsible for the social aspects of the referential use of language: it is a switch which takes reference as internal into our brains. What do change, however, are the methodological and the "metaphysical" presuppositions. No strict semantic equivalences between the senses (expressed by descriptions) and proper names are required; no unscrutinized intuitions

are accepted as evidence, so the "one-shot counterexample" strategy, so dear to philosophers, is rejected. Reference is seen as a form of neurocognitive information-based action through language use, not an abstract picture, a mere (causal) relation between a term and an object. These are manifestations of the typical effects of the kinds of general changes in the interpretation of phenomena that characterize theoretical *Gestalt* switches.

The results of this particular switch seem also to be characteristic to such switches, though they may not so seem from the start. I dare to claim that for the first time now we have the beginnings of a unifying framework under which the diverse phenomena can be subsumed. For instance the distinction genuine reference-purported reference is seen to be not so important explanatorily (recall the case of the empty names). Important old problems retain in essence their sense theoretic explanations (the "four puzzles"). Some problems are seen not to be such at all, for instance the counterexamples proposed by the direct-causal theorists fail, and their *prima facie* force can be explained by the basic schema of reference, the practice of deference and the epistemic symmetry principle. And finally, what is left of the direct-causal theory of reference becomes, or rather is seen to have always been, part of the Fregean sense theory of reference. The causal relations from the referents are informational and as such have never been in dispute by the sense theorists. The real question is: what kinds of causal relations are referential? The answer the neuronaturalized sense theory of reference delivers is: the neurobiologically based perceptuo-cognitive information processings (with motor and emotional informational elements). These internal schema representations conceived as relations are of course also causal, in the common way of speaking, but it is their informativity that counts explanatorily, as it counts with the external causal relations outside one's brain.

Moreover our "inside the skull" view has in fact overwhelming evidence in its favour. We have been staring at it all the time. It goes on automatically, and only occasionally is there an error (an incorrect description or piece of information). This is the point I raised earlier that the senses expressed by the descriptions have been serving us so well as the vehicles of reference that the concocted and rare types of cases of alleged reference failure the direct-causal theorists have imagined to be relevant can be put aside. We communicate and refer to individuals, events, natural and technological kinds, concepts and so on fluently as if it were our second nature. Instead it is our *first* nature, I should think, as the cognitive animals that we are. The naturalness of my framework should be seen against this background. We live *in* it, in the continuous pulses of the tremendous amounts of neurocognitive information in the intentional-cum-biological acts, and because of that also in the referential acts, in the one essential element of our neuronaturalistic *Dasein*.

Notes

1. Though there are "deviant" uses of these expressions, for example when one can refer to a certain pudding years after eating it by saying "that plum pudding was delicious".

BIBLIOGRAPHY

- Adams, R. 1979: "Primitive Thisness and Primitive Identity", *Journal of Philosophy* 76, 5-26.
- Almog, J. 1986: "Naming Without Necessity", *Journal of Philosophy* 83, 210-42.
- Almog, J. 2005: "Is a Unified Description of Language-and-Thought Possible?", *Journal of Philosophy* 102, 493-531.
- Almog, J., Perry, J. & Wettstein, H. (eds.): *Themes from Kaplan*, New York, Oxford University Press.
- Arbib, M. & Hesse, M.: 1986: *The Construction of Reality*, Cambridge, Cambridge University Press.
- Bach, K. 1987: *Thought and Reference*, Oxford, Clarendon Press.
- Bar-Elli, G. 1981: "Frege and the Determination of Reference", *Erkenntnis* 16, 137-60.
- Beaney, M. 1997: *The Frege Reader*, Oxford, Blackwell Publishers.
- Bell, D.: 1984: "Reference and Sense: An Epitome", in C. Wright (ed.): *Frege: Tradition & Influence*, Oxford, Basil Blackwell, 184-8.
- Bell, D. 1990: "How Russellian was Frege?", *Mind* 99, 167-77.
- Berger, A. 1989: "A Theory of Reference Transmission and Reference Change", in P. French *et al.* (eds.), 180-98.
- Berger, A. 2002: *Terms and Truth. Reference Direct and Anaphora*, Cambridge, MA., MIT Press.
- Bower, G.: 1981: "Mood and Memory", *The American Psychologist* 36, 283-300.
- Brock, S. 2004: "The Ubiquitous Problem of Empty Names", *Journal of Philosophy* 101, 277-98.
- Brothers, L. 1990: "The Social Brain: A Project for Integrating Primate Behavioral Neurophysiology in a New Domain", *Concepts in Neuroscience* 1, 27-51.
- Buck, R. 1987: "The Psychology of Emotion", in J. LeDoux & W. Hirst (eds.), 275-300.
- Burge, T. 1977. "Belief *De Re*", *Journal of Philosophy* 74, 338-62.
- Burge, T. 1979: "Sinning Against Frege", *Philosophical Review* 88, 398-432.
- Butterworth, G. & Glover, L. 1988: "Origins of Referential Communication", in L. Weiskrantz (ed.): *Thought Without Language*, Oxford, Clarendon Press, 5-24.
- Cappelen, H. & Winblad, D. 1999: "Reference Externalized and the Role of Intuitions in Semantic Theory", *American Philosophical Quarterly* 36, 337-50.
- Caramazza, A. & McCloskey, M. 1988: "The Case for Single-Patient Studies", *Cognitive Neuropsychology* 5, 517-28.
- Castaneda, H.-N. 1981a: "On the Philosophical Foundations of the Theory of Communication: Reference", in P. French *et al.*, 125-46.
- Castaneda, H.-N. 1981b: "The Causal and the Epistemic Roles of Proper names in Our Thinking of Particulars", in P. French *et al.*, 151-8.
- Chastain, C. 1975: "Reference and Context", in K. Gunderson (ed.): *Minnesota Studies in the Philosophy of Science, VII, Language, Mind and Knowledge*, Minneapolis, University of Minnesota Press, 194-269.
- Conway, M. (ed.) 1997: *Cognitive Models of Memory*, Hove, Psychology Publishers.
- Damasio, A. 1999: *The Feeling of What Happens. Body and Emotion in the Making of Consciousness*, New York, Harcourt, Brace & Company.
- Damasio, A. & Damasio, H. 1993: "Brain and Language", in *Mind and Brain. Readings in Scientific American*, New York, W. H. Freeman and Company, 54-65.
- Davidson, D. & Harman, G. (eds.): *Semantics of Natural Language*, Dordrecht, D. Reidel.
- Davies, M. & Humberstone, L. 1980: "Two Notions of Necessity", *Philosophical Studies* 38, 1-30.
- Devitt, M. 1974: "Singular Terms", *Journal of Philosophy* 71, 183-205.

- Devitt, M. 1981: *Designation*, New York, Columbia University Press.
- Devitt, M. 1989: "Against Direct Reference", in P. French *et al.* (eds.), 206-40.
- Devitt, M. 1994: "The Methodology of Naturalistic Semantics", *Journal of Philosophy* 91, 545-72.
- Devitt, M. & Sterelny, K. 1987: *Language and Reality. An Introduction to the Philosophy of Language*, Oxford, Basil Blackwell.
- Donnellan, K. 1972: "Proper names and Identifying Descriptions", in D. Davidson & G. Harman, 356-79.
- Donnellan, K. 1974: "Speaking of Nothing", *Philosophical Review* 83, 3-31.
- Donnellan, K. 1979: "The Contingent *A Priori* and Rigid Designators", in P. French *et al.* (eds.): *Contemporary Perspectives in the Philosophy of Language*, Minneapolis, University of Minnesota Press.
- Dummett, M. 1973: *Frege: Philosophy of Language*, London, Duckworth.
- Dummett, M. 1981: *The Interpretation of Frege's Philosophy*, Cambridge, MA., Harvard University Press.
- Edelman, G. 1987: *Neural Darwinism. The Theory of Neuronal Group Selection*, New York, Basic Books, Inc.
- Edelman, G. 1989: *The Remembered Present. A Biological Theory of Consciousness*, New York, Basic Books, Inc.
- Ellis, A. & Young, A. 1988: *Human Cognitive Neuropsychology*, Hove, Lawrence Erlbaum Associates.
- Ellis, A. & Young, A. 1998: "Accounting for Delusional Misidentification", in A. Young (ed.), 125-44.
- Evans, G. 1977: "The Causal Theory of Names", in S. Schwartz (ed.), 192-215.
- Evans, G. 1982: *The Varieties of Reference*, J. McDowell (ed.), Oxford, Clarendon Press.
- Evans, G. 1985: "Pronouns, Quantifiers and Relative Clauses (I)", in *Collected Papers*, Oxford, Clarendon Press, 76-152.
- Feyerabend, P. 1981: "Materialism and the Mind-Body Problem" in *Realism, Rationalism and Scientific Method. Philosophical Papers, Volume I*, Cambridge, Cambridge University Press, 161-75.
- Field, H. 1973: "Theory Change and the Indeterminacy of Reference", *Journal of Philosophy* 70, 462-81.
- Fischbein, E. 1987: *Intuition in Science and Mathematics. An Educational Approach*, Dordrecht, D. Reidel.
- Fodor, J. 1987: *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*, Cambridge, MA., MIT Press.
- Fodor, J. & Lepore, E. 1996a: "Paul Churchland and State Space Semantics", in R. McCauley (ed.), 145-59.
- Fodor, J. & Lepore, E. 1996b: "Reply to Churchland", in R. McCauley (ed.), 159-62.
- Fodor, J. & Lepore, E. 1999: "All at Sea in Semantic Space: Churchland on Meaning Similarity", *Journal of Philosophy* 96, 381-403.
- Forbes, G. 1987: "Indexicals and Intensionality: A Fregean Perspective", *Philosophical Review* 96, 3-33.
- Forbes, G. 1989: *Languages of Possibility. An Essay in Philosophical Logic*, Aristotelian Society Series, Volume 9, London, Basil Blackwell.
- Forbes, G. 1989: "Cognitive Architecture and the Semantics of Belief", in P. French *et al.* (eds.), 84-100.
- Forbes, G. 1990: "The Indispensability of Sinn", *Philosophical Review* 99, 536-63.
- Forbes, G. 1997: "Essentialism", in B. Hale & C. Wright (eds.).
- Freeman, W. 1999: *How Brains Make Up Their Minds*, London, Weidenfeld & Nicolson.

- Frege, G. 1997a: "On *Sinn* and *Bedeutung*", in M. Beaney, 151-71.
- Frege, G. 1997b: "Thought", in M. Beaney, 325-45.
- Frege, G. 1997c: "On Concept and Object", in M. Beaney, 181-93.
- Frege, G. 1997d: "Letter to Husserl, 24.5.1891", in M. Beaney, 149-50.
- Frege, G. 1997e: "Comments on *Sinn* and *Bedeutung*", in M. Beaney, 172-80.
- Frege, G. 1997f: "Letter to Jourdain, Jan. 1914", in M. Beaney, 319-21.
- French, P., Uehling, T. & Wettstein, H. (eds.) 1981: *Midwest Studies in Philosophy. Contemporary Perspectives on the Philosophy of Language*, 2nd printing, Minneapolis, University of Minnesota Press.
- French, P., Uehling, T. & Wettstein, H. (eds.) 1989: *Midwest Studies in Philosophy. Contemporary Perspectives in the Philosophy of Language*, Notre Dame, Notre Dame University Press.
- Garnham, A. 1997: "Representing Information in Mental Models", in M. Conway (ed.), 149-76.
- Gawron, J. & Peters, S. 1990: *Anaphora and Quantification in Situation Semantics*, Stanford, CSCI Lecture Notes, Number 19.
- Gazzaniga, M. (ed.) 1995: *The Cognitive Neurosciences*, Cambridge, MA., MIT Press.
- Georgopoulos, A., Lurito, J., Petrides, M., Schwartz, A. & Massey, J. 1989: "Mental Rotation of the Neuronal Population Vector", *Science* 24, 234-6.
- Gibson, J. 1966: *Senses Considered as Perceptual Systems*, Boston, Houghton Mifflin.
- Hale, B. & Wright, C. (eds.) 1997: *Blackwell Companion to Philosophy. A Companion to the Philosophy of Language*, Oxford, Blackwell Publishers.
- Hingley, R. 1974: *Joseph Stalin: Man and Legend*, London, Hutchinson & Co.
- Hintikka, J. 1969: *Models for Modalities. Selected Essays*, Dordrecht, D. Reidel.
- Hintikka, J. 1969a: "Semantics for Propositional Attitudes", in Hintikka 1969, 87-111.
- Hintikka, J. 1975: *Intentions of Intentionality and Other New Models for Modalities*, Dordrecht, D. Reidel.
- Hintikka, J. 1975a: "The Semantics of Modal Notions and the Indeterminacy of Ontology", in J. Hintikka 1975, 26-42.
- Hintikka, J. 1975b: "Different Constructions in Terms of the Base Epistemic Verbs", in J. Hintikka 1975, 1-25.
- Hintikka, J. 1975c: "Carnap's Heritage in Logical Semantics", in J. Hintikka 1975, 76-101.
- Hintikka, J. 1989: "The Cartesian *Cogito*, Epistemic Logic and Neuroscience", in J. Hintikka: *The Logic of Epistemology and the Epistemology of Logic*, Dordrecht, D. Reidel, 113-36.
- Hintikka, J. 1999: "The Emperor's New Intuitions", *Journal of Philosophy* 96, 127-47.
- Hintikka, J. & Sandu, G. 1995: "The Fallacies of the New Theory of Reference", *Synthese* 104, 245-83.
- Hintikka, J. & Symons, J. 2003: "Systems of Visual Identification in Neuroscience: Lessons from Epistemic Logic", *Philosophy of Science* 70, 89-104.
- Holloway, R. 1996a: "Evolution of the Human Brain", in A. Lock & C. Peters (eds.), 74-108.
- Holloway, R. 1996b: "Editorial appendix II: Evolution of the Human Vocal Apparatus", in A. Lock & C. Peters (eds.), 116-25.
- Jackson, F. 2004: "Why We Need A-Intensions", *Philosophical Studies* 118, 257-77.
- Kandel, E. & Schwartz, J. (eds.) 1985: *Principles of Neural Science*, 6th printing, New York, Elsevier Science Publishing Company, Inc.
- Kaplan, D. 1989: "Afterthoughts", in J. Almog *et al.* (eds.), 565-614.
- Kaplan, D. 1990: "Dthat", in P. Yourgrau (ed.): *Demonstratives*, Oxford, Oxford University Press, 11-33.
- Knight, R. & Grabowcky, M. 1995: "Escape from Linear Time: Prefrontal Cortex and Conscious Experience", in M. Gazzaniga (ed.), 1357-71.

- Kripke, S. 1972: "Naming and Necessity", in D. Davidson & G. Harman (eds.), 253-355.
- Kripke, S. 1977: "Identity and Necessity", in S. Schwartz (ed.), 66-101.
- Kripke, S. 1979: "A Puzzle About Belief", in A. Margalit (ed.): *Meaning and Use*, Dordrecht, D. Reidel.
- Kripke, S. 1980: *Naming and Necessity*, Oxford, Basil Blackwell.
- Kutas, M. & Dale, A. 1997: "Electrical and Magnetic Recordings of Mental Functions", in M. Rugg (ed.), 197-242.
- Kvart, I. 1989: "Divided Reference", in P. French *et al.*, 140-79.
- Larkman, A. & Jack, J.: 1995: "Synaptic Plasticity: Hippocampal LTP", *Current Opinion in Neurobiology* 5, 324-34.
- LeDoux, J. & Hirst, W. (eds.): *Mind and Brain. Dialogues in Cognitive Neuroscience*, New York, Cambridge University Press.
- LeDoux, J. 1987: "The Neurobiology of Emotion", in J. LeDoux & W. Hirst (eds.), 301-54.
- Linsky, L. 1967: *Referring*, London, Routledge and Kegan Paul.
- Linsky, L. 1977: *Names and Descriptions*, Chicago, The University of Chicago Press.
- Llinas, R. & Churchland, P. S. (eds.) 1996: *Mind-Brain Continuum: Sensory processes*, Cambridge, MA., MIT Press.
- Llinas, R. & Paré, D. 1996: "Brain as a Closed System Modulated by the Senses", in R. Llinas & P. S. Churchland (eds.), 1-18.
- Loar, B. 1976: "The Semantics of Singular Terms", *Philosophical Studies* 30, 353-77.
- Lock, A. & Peters, C. (eds.): *Handbook of Human Symbolic Evolution*, Oxford, Clarendon Press.
- Luntley, M. 1999: *Contemporary Philosophy of Thought. Truth, World, Content*, Oxford, Blackwell Publishers.
- Mainzer, K. 1996: *Symmetries of Nature. A Handbook for Philosophy of Nature and Science*, Berlin, Walter de Gruyter.
- Makin, G. 2000: *The Metaphysics of Meaning. Russell and Frege on Sense and Denotation*, London, Routledge.
- Marcus, R. 1961: "Modalities and Intensional Languages", *Synthese* 13, 303-22.
- Martin, J. 1985: "Cortical Neurons, the EEG, and the Mechanisms of Epilepsy", in E. Kandel & J. Schwartz (eds.), 636-47.
- Marzke, M. 1996: "Evolution of Hand and Bipedality", in A. Lock & C. Peters (eds.), 126-54.
- Maunsell, J. & Ferrera, V. 1995: "Attentional Mechanisms in Visual Cortex", in M. Gazzaniga, 451-61.
- May, R. 2006: "The Invariance of Sense", *Journal of Philosophy* 103, 111-44.
- McCarthy, R. 1993: "Assembling Routines and Addressing Representations: An Alternative Conceptualization of 'What' and 'Where' in the Human Brain", in N. Eilan, R. Mc Carthy & B. Brewen (eds.): *Spatial Representation. Problems in Philosophy and Psychology*, Oxford, Blackwell, 373-99.
- McCauley, R. (ed.) 1996: *The Churchlands and Their Critics*, Oxford, Blackwell Publishers Ltd.
- McDowell, J. 1977: "On the Sense and Reference of a Proper Name", *Mind* 86, 159-85.
- McKinsey, M. 1978: "Names and Intentionality", *Philosophical Review* 87, 174-200.
- Messier, D. & Collis, G. 1996: "Early Interaction and Cognitive Skills: Implications for the Acquisition of Culture", in A. Lock & C. Peters (eds.), 432-68.
- Mill, J.S. 1961: *A System of Logic*, 8th ed., London, Longmans.
- Mountcastle, V. 1998: *Perceptual Neuroscience. The Cerebral Cortex*, Cambridge, MA., Harvard University Press.
- Neisser, U. 1976: *Cognition and Reality. Principles and Implications of Cognitive Psychology*, San Francisco, Freeman.

- Nelson, J. 1992: *Naming and Reference. The Link of Word to Object*, London and New York, Routledge.
- Noonan, H. 2001: *Frege. A Critical Introduction*, Cambridge, Polity.
- Pandya, D. & Yeterian, E. 1990: "Architecture and Connections of Cerebral Cortex: Implications for Brain Evolution and Function", in A. Scheibel & A. Wechsler (eds.), 53-84.
- Pap, A. 1958: *Semantics and Necessary Truth. An Inquiry into the Foundations of Analytical Philosophy*, New Haven, Yale University Press.
- Parkin, A. 1996: *Explorations in Cognitive Neuropsychology*, Oxford, Blackwell.
- Peregrin, J. 2000: "Reference and Inference: The Case of Anaphora", in K. van Heusinger & U. Egli (eds.): *Reference and Anaphoric Relations*, Dordrecht, D. Reidel, 269-86.
- Phillips, W. 1997: "Theories of Cortical Computation", in M. Rugg (ed.), 11-46.
- Phillips, W. & Singer, W. 1997: "In Search of Common Foundations for Cortical Computation", *Behavioral and Brain Sciences* 20, 657-722.
- Plantinga, A. 1978: "The Boethian Compromise", *American Philosophical Quarterly* 15, 129-38.
- Putnam, H. 1986a: *Reason, Truth and History*, reprint, Cambridge, Cambridge University Press.
- Putnam, H. 1986b: *Realism and Reason. Philosophical Papers, Vol. 3*, Cambridge, Cambridge University Press.
- Putnam, H. 1987: *Mind, Language and Reality. Philosophical Papers, Vol.2*, Cambridge, Cambridge University Press.
- Putnam, H. 1987a: "Is Semantics Possible?", in H. Putnam 1987, 139-52.
- Putnam, H. 1987b: "Explanation and reference", in H. Putnam 1987, 196-214.
- Putnam, H. 1987c: "The Meaning of 'Meaning'", in H. Putnam 1987, 215-71.
- Putnam, H. 1988: *Representation and Reality*, Cambridge, MA., MIT Press.
- Quine, W. 1988: *Word & Object*, Cambridge, MA., MIT Press.
- Recanati, F. 1988: "Rigidity and Direct Reference", *Philosophical Studies* 53, 103-17.
- Recanati, F. 1989: "Referential/Attributive: A Contextualist Proposal", *Philosophical Studies* 56, 217-49.
- Recanati, F. 1993: *Direct Reference. From Language to Thought*, Oxford, Blackwell Publishers.
- Recanati, F. 2000: *Oratio Obliqua, Oratio Recta. Analysis in Metarepresentation*, Cambridge, MA., MIT Press.
- Rizzolatti, G. & Arbib, M. 1998: "Language Within Our Grasp", *Trends in Neuroscience* 21, 188-94.
- Robin, N. & Holyoak, K. 1995: "Relational Complexity and the Functions of Prefrontal Cortex", in M. Gazzaniga (ed.), 987-97.
- Rolls, E. & Tovéé, M. 1994: "Processing Speed in the Cerebral Cortex and the Neurophysiology of Visual Masking", *Proceedings of the Royal Society of London, series B*, 257, 9-15.
- Rosenberg, J. 1994: *Beyond Formalism. Naming and Necessity for Human Beings*, Philadelphia, Temple University Press.
- Rugg, M. (ed.) 1997: *Cognitive Neuroscience*, Hove, Psychology Press, Publishers.
- Russell, B. 1905: "On Denoting", *Mind* 14, 479-93.
- Russell, B. 1912: *The Problems of Philosophy*, Oxford, Oxford University Press.
- Russell, B. 1948: *Human Knowledge. Its Scope and Limits*, London, George Allen and Unwin Ltd.
- Russell, B. 1956: "The Philosophy of Logical Atomism" in C. Marsh (ed.): *Logic and Knowledge: Essays 1901-1950*, London, George Allen and Unwin, 177-281.

- Russell, B. 1993: "Descriptions" in A. Moore (ed.): *Meaning and Reference*, Oxford, Oxford University Press, 46-55.
- Salmon, N. 1979: "Names and Descriptions", *Journal of Philosophy* 76, 436-52.
- Salmon, N. 1982: *Reference and Essence*, Oxford, Basil Blackwell.
- Salmon, N. 1986: *Frege's Puzzle*, Cambridge, MA., MIT Press.
- Salmon, N. 1989: "Reference and Information Content: Names and Descriptions", in D. Gabbay & F. Guentner (eds.): *Handbook of Philosophical Logic. Vol. IV: Topics in the Philosophy of Language*, Dordrecht, D. Reidel, 409-62.
- Sargent, J., MacDonald, B. & Zuck, E. 1994: "Structural and Functional Organization of Knowledge About Faces and Proper names: A Positron Emission Tomography Study", in C. Umiltá & M. Moscovitch (eds.), 203-28.
- Sbisá, M. 2003: "Foreword", *Facta Philosophica* 5, 155-61.
- Scheibel, A. & Wechsler, A. (eds.) 1990: *Neurobiology of Higher Cognitive Function*, New York, Guilford Press.
- Schiffer, S. 1978: "The Basis of Reference", *Erkenntnis* 13, 171-206.
- Schiffer, S. 2005: "Russell's Theory of Definite Descriptions", *Mind* 114, 1135-83.
- Schwartz, S. (ed.) 1977: *Naming, Necessity and Natural Kinds*, Ithaca, New York, Cornell University Press.
- Searle, J. 1958: "Proper names", *Mind* 66, 166-73.
- Searle, J. 1967: "Proper names and Descriptions" in P. Edwards (ed.): *The Encyclopedia of Philosophy*, 6, New York, 487-92.
- Semenza, C. & Zettin, M. 1988: "Generating Proper names: A Case of Selective Inability", *Cognitive Neuropsychology* 5, 711-21.
- Shanks, D. 1997: "Representation of Categories and Concepts in Memory", in M. Conway (ed.), 111-46.
- Shapiro, M. & Eichenbaum, H. 1997: "Learning and Memory: Computational Principles and Neural Mechanisms", in M. Rugg (ed.), 77-130.
- Shastri, L. & Ajjanagadde, V. 1993: "From Simple Associations to Systematic Reasoning: A Connectionist Representation of Rules, Variables and Dynamic Binding with Temporal Synchrony", *Behavioral and Brain sciences* 16, 417-94.
- Simmons, J. 1996: "Formation of Perceptual Objects from the Timing of Neural Responses: Target-Range Images on Bat Sonar", in R. Llinas & P. S. Churchland (eds.), 219-50.
- Singer, W. 1996: "Neuronal Synchronization: A Solution to the Binding Problem", in R. Llinas & P. S. Churchland (eds.), 101-30.
- Smith, A. 1984: "Rigidity and Scope", *Mind* 93, 177-93.
- Smith, Q. 1995a: "Marcus, Kripke, and the Origin of the New Theory of Reference", *Synthese* 104, 179-89.
- Smith, Q. 1995b: "Marcus and the New Theory of Reference: A Reply to Scott Soames", *Synthese* 104, 217-44.
- Soames, S. 1995: "Revisionism About Reference: A Reply to Smith", *Synthese* 104, 192-210.
- Soames, S. 2002: *Beyond Rigidity. The Unfinished Semantic Agenda of 'Naming and Necessity'*, New York, Oxford University Press.
- Soames, S. 2005: *Reference and Description. The Case Against the Two-Dimensionalism*, Princeton, Princeton University Press.
- Spelke, E., Vishton, P. & van Hofsten, C. 1995: "Object Perception, Object-Directed action, and Physical Knowledge in Infancy", in M. Gazzaniga (ed.), 165-79.
- Stanley, J. 1997: "Names and Rigid Designation", in B. Hale & C. Wright (eds.), 555-85.
- Stich, S. 1996: *Deconstructing the Mind*, New York, Oxford University Press.
- Strawson, P. 1959: *Individuals. An Essay in Descriptive Metaphysics*, London, Methuen & Co Ltd.

- Strawson, P. 1971: *Logico-Linguistic Papers*, London, Methuen & Co Ltd.
- Strawson, P. 1971a: "Identifying Reference and Truth-Values" in P. Strawson 1971, 75-95.
- Strawson, P. 1971b: "On Referring" in P. Strawson 1971, 1-27.
- Tononi, G., Sporns, O. & Edelman, G. 1992: "Re-entry and the Problem of Integrating Multiple Cortical Areas: Simulation of Dynamic Interaction in the Visual System", *Cerebral Cortex* 2, 310-35.
- Tononi, G. & Edelman, G. 1997: "Information: In the Stimulus or in the Context?", *Behavioral and Brain Sciences* 20, 698-700.
- Tranel, D. & Damasio, A. 1993: "The Covert Learning of Affective Valence Does Not Require Structures in Hippocampal System or Amygdala", *Journal of Cognitive Neuroscience* 5, 78-88.
- Treisman, A. 1996: "The Binding Problem", *Current Opinion in Neurobiology* 6, 171-8.
- Treves, A. & Rolls, E. 1994: "Computational Analysis of the Role of Hippocampus in Memory", *Hippocampus* 4, 374-91.
- Umlá, C. & Moscovitch, M. (eds.): *Attention and Performance XV. Conscious and Nonconscious Information Processing*, Cambridge, MA., MIT Press.
- Wettstein, H. 1981: "Proper names and Propositional Opacity", in P. French *et al.*, 147-50.
- Wettstein, H. 1986: "Has Semantics Rested on a Mistake?", *Journal of Philosophy* 83, 185-209.
- Wettstein, H. 1991: *Has Semantics Rested on a Mistake? And Other Essays*, Stanford, Stanford University Press.
- Wettstein, H. 2004: *The Magic Prism. An Essay in the Philosophy of Language*, Oxford, Oxford University Press.
- Wynn, T. 1996: "The Evolution of Tools and Symbolic Behavior", in A. Lock & C. Peters (eds.), 263-87.
- Yin, J. & Tully, T. 1998: "CREB and the Formation of Long-Term Memory", *Current Opinion in Neurobiology* 6, 264-8.
- Young, A. (ed.) 1998: *Face and Mind*, Oxford, Oxford University Press.
- Yu, P. 1980: "The Modal Argument Against the Description Theories of Names", *Analysis* 40, 208-9.

