A User's Guide to Proper Names

Names
Their Pragmatics And Semantics

Anna Pilatova

A User's Guide to Proper Names

Their Pragmatics And Semantics

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Rijksuniversiteit Groningen

A User's Guide to Proper Names
Their Pragmatics And Semantics

PROEFSCHRIFT

ter verkrijging van het doctoraat in de
Wijsbegeerte
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts
in het openbaar te verdedigen op
donderdag 8 september 2005
om 13.15 uur

door

Anna Pilatova

geboren op 30 maart 1971
te Praag (Tsjechië)

To my grandmothers, Rachil and Jaryna,

some of the strongest and most extraordinary women

I have had the honour to meet.

Contents

Acknowledgments

I would like to thank everyone who helped me to write this thesis, and apologise to anyone whom I might forget to mention. I received support, ideas, and encouragement from institutions, supervisors, colleagues, my students, people at conferences, my friends, and family. The writing of this thesis co-incided with the years I spent in The Netherlands, and I want to use this opportunity to express my gratitude to the people who were during this time a part of my life, who helped me make Amsterdam my home, and Groningen my second home. Hence the unusual length of this section.

While writing this thesis, I was a part of two departments, and had the good fortune of being able to profit from many events and lectures organised both by the Department of Philosophy of the University of Groningen and by the Department of Philosophy of the University of Amsterdam, and the ILLC, with which this department is associated. My research was supported by a grant I received from the Ubbo Emmius Funds, which are administered by the University of Groningen. I want to thank Prof. Theo Kuipers, who allowed me to divide my time between Groningen and Amsterdam. Without Steve Thomas, I would have probably never moved to Amsterdam in the first place. I want to thank him for his friendship and support.

I was extraordinarily fortunate in having Erik Krabbe and Martin Stokhof as my supervisors. I could not have finished this project without their professional and personal support or without their

Katherine Gardiner's home was my home as well. I saw Joran, her son, grow from a baby into a young schoolboy, and Nynke, her daughter, transform from a child into a self-possessed young lady. It has been a privilege. I want to thank both Nynke and Joran for their love. Katherine's friendship supported me through many a difficult time, and there is no way I can thank her enough. She and her family will always have a special place in my heart.

I was very fortunate in my flatmates: thanks to Andy Liu, European intellectual *par excellence*, from whom I learned so much about China, its people and history, thanks to Mehmet Altinakya, he knows why, thanks to Veronica Steward for her liveliness and lovely Scottish accent, thanks to Marie Safarova, who helped me through a difficult time, may we never lose touch, thanks to Alastair Butler, who introduced me to LaTex and British virtues, and convinced me that one can exclusively on potatoes, carrots, and corned chicken or quiche alone. Thanks also to Elise, now Alastair's wife, for her friendship and trust. The final stages of the thesis were written at a hectic pace, and I took to working through the nights. As a consequence, my flatmates Inga Bauer and Tim Scheufen saw me mainly just in my pyjamas. They were very tolerant and kind.

Henk Zeevat and his friendship has been invaluable. Our arguments about philosophy, history, literature, the advice on DIY and bicycle maintenance – thank you so much for everything. The love and friendship of Russell Dale has been with me for little less than half of my life, proving that though it is difficult, it is possible to stay in touch mainly by correspondence.

Since 2002, when the Heideggerian Hermeneutic Club came into existence, it has been a place for me to relax, connect with people, and get out at least once a week. I want to thank all of its members, present and past. That is, its founding members, David Wood, the president, Fabrice Nauze, the co-president and secretary for emancipation, Will Rose, always ready to argue with Fabrice, Marian Counihan, the resident South African. Some members who joined a little later also gained prominence: Marie Safarova, who introduced Czech as a second official language of the club, Gilad Mishne and Oren Tsur, who helped me form the Jewish faction, Albin Gi-

rardin, always a gentleman, thanks for your unforgettable singing, Helle Hansen, with her sound common sense and adventurous ideas, David Ahn, who single-handedly overturned my prejudice that men in suits are boring, Lotta Weckstrom, who gave us at one point a rare insight into the life of people who are not in the academia, Samson de Jager, always sound yet fun, a resident New Zealander, Breanndan O'Nuallain, an associated member, a representative of the Irish faction. I miss you all.

I want to thank Oren Schwartz for making me take some time for culture and snooker in the summer of 2004. Simone Blayer and his family fed me in the last difficult stages of finishing my thesis. Their company has always been a safe harbour to me. Thank you for your friendship.

Alastair Butler taught me the basics of using LaTex. After he left for Singapore, Samson de Jager helped me out on numerous occassions. Cheers. I also want to thank Mary Lommerse, who helped me keep sane through my professional and personal crises, and all my friends from the Yiddish choir Heimish Zeyn, who made me feel heimish indeed.

And last but not least: my family. My mother always believed implicitly that academically, I could do whatever I put my mind to. Her mother, my grandmother, always believed I had a good common sense, and could basically deal with anything. From my father I think I got the sense of humour allowing me to cope somehow with these expectations. How could I thank you enough? While I was writing my thesis, my brother found he, too, likes the academic environment, and I wish him best of luck with his thesis. My twin sister has been taking me for holidays to exotic places so often my friends called me a 'jet-setter'. With her, even a rainy day in Amsterdam could be exciting.

Amsterdam   Anna Pilatova
June, 2005.

Preface

The origins of this work go all the way to my reading of Kripke's *Naming and Necessity* in 1993. It had left me with a feeling of dissatisfaction that lingered long enough to inspire my MA thesis (on *Internalism and Externalism in the Theories of Reference*), and finally inspired the current work. Over time, I became acquainted with other influential accounts of reference of proper names, but my unease with essentialism and wariness of direct designation remained.

In the *User's Guide to Proper Names*, I seized the opportunity to analyse what I thought was problematic about the mainstream theories of reference of proper names. I tried to tease apart a number of related doctrines about the behaviour of proper names, and thus arrive at a better understanding of how the various parts of the theories of reference I chose to analyse are related. This helped me to develop my own proposal regarding both the semantics and the pragmatics of proper names.

The *Guide* is organised around a particular approach to the tasks of a semantic theory. According to this approach, proposed by Stalnaker, a semantic theory dealing with proper names should account for the descriptive semantics of names, their foundational semantics, and the semantics of modal statements in which they figure. Descriptive semantics focuses on the contribution a proper name makes to the truth-value of sentences in which it occurs. Based on such an analysis, a proper name is assigned a semantic value, which

is supposed to provide us with an interpretation of that name. A crucial part of this task is to see just what kind of thing the semantic value of a proper name is.

In interpreting sentences containing proper names, one can, and often does, use the notion of possible worlds. This is especially true if the sentences in question are modal. There are various approaches to modality, which carry with them different sets of presuppositions. An analysis of the systematic features and presuppositions of various possible-world frameworks is a part of the task of the semantics of modal statements. Another task of semantics of modal statements is to investigate where the constraints on possible worlds used in analysing these statements derive from, that is, whether, and to what degree, they should derive from the descriptive or the foundational part of the semantic enterprise.

Both of these parts of a semantic theory have the potential to make predictions about the foundational semantics of proper names, which deals with the speaker's behaviour and communication. Foundational semantics aims at answering the following question: What makes it the case that the language spoken by a particular individual or community has the very descriptive semantics it has? In this investigation, one looks at the speaker, her intentions and communicative goals, and tries to identify the strategies she uses to get her (linguistic) point across.

The *User's Guide to Proper Names* is divided in two basic parts. The first part, encompassing chapters one through to four, is devoted to a reconstruction and analysis of several influential approaches to the semantics of proper names. The second part of the thesis, chapters five to seven, contains my own proposal regarding the semantics of proper names and a conclusion.

The first part of the thesis starts with an introductory chapter, *Outlining the Field and Introducing Some of the Players*. Its task is to show the usefulness of organising the thesis around the distinction between descriptive, modal, and foundational semantics, and to provide a proper characterisation of these notions. In the second part of the same chapter, *Setting the Scene*, I give a preliminary outline of some of the notions needed to describe the theories of reference

that are later investigated, that is mainly the theories proposed by Kripke, Kaplan, Lewis, and Stalnaker.

The second chapter, *Descriptive Semantics*, focuses on the descriptive semantics proposed by Kripke and Kaplan for proper names. In the first part of the chapter, I introduce several arguments in favour of rigid designation. A closer look reveals that none of the three arguments in question – the modal, the epistemological, and the semantic one – is actually an argument *for* rigid designation. They all just argue *against* some forms of descriptivism. Moreover, in the case of the modal and the epistemological argument, it is relatively easy to find forms of descriptivism that are immune to the lines of reasoning proposed in the arguments. The semantic argument seems to be the strongest of the three because it relies on a straightforward clash of intuitions regarding the identity of individuals.

An analysis of a direct argument for rigid designation also highlights the connection between rigid designation and certain preconceptions about the identity of individuals. A further investigation of issues related to the identity of individuals across possible worlds emerges at this point as an important issue to tackle. It also becomes clear that rigid designation alone cannot fully determine the Kripkean picture of names as non-descriptive entities referring without a mediation of any sort of conceptual content.

In the second part of Chapter 2, I introduce the basic notions of Kaplan's approach to the descriptive semantics of proper names. It turns out that in order to derive the familiar Kripkean picture of proper names, one has to presuppose direct reference for names and at least some version of haecceitism for the individuals in question. While, as we show, Kaplan's framework does not work well for proper names, it gives us some conceptual tools that help us in our undertaking.

In the third chapter, *Modal Statements, Individuals, and Essences*, I analyse three different approaches to building a possible-world framework, Lewis's, Kripke's, and Stalnaker's. In each case, the same questions are asked: What is the motivation and intended field of application of this framework? What are the ontological

commitments of the approach? What form of essentialism, if any, does it imply? How does it deal with the notion of an individual?

In Lewis's case, the main problem turns out to be the theory of counterparts, which is, as I show, an integral part of his approach. The concept of an individual implied by it does not seem to correspond to any intuitive reading of counterfactual statements involving individuals. The investigation of Kripke's framework focuses on describing the weakest form of essentialism that has to be presupposed to make the proposal work. Once that is concluded, I analyse the essentialism Kripke actually proposes, and the motivation and presuppositions on which it rests. I conclude that its motivation cannot be said to come from an analysis of language and that it presupposes a particular form of scientific realism. The rest of the chapter is devoted to a reconstruction and analysis of Stalnaker's possible world framework, which turns out to be rather more cautious about metaphysical presuppositions and better suited for an analysis of natural language. As in the two previous proposals, I try to reconstruct the notion of actual world that is presupposed here.

In the fourth chapter, *Foundational Semantics*, I investigate the notion of proposition implied by Lewis's, Kripke's, and Stalnaker's approach. I focus on propositions containing proper names, and analyse the way in which each of the conceptions mentioned above is vulnerable to the problem of logical omniscience. An analysis of Lewis's framework reveals that the concept of proposition implied by it is so weak as to be rather uninteresting. An investigation of Kripke's concept of proposition deals not only with the systematic issues, but also with the Pierre puzzle and various attempts at solving it. I present a number of different approaches to the problem and compare their merits. Stalnaker is very worried about the problem of logical omniscience, and yet it turns out that his conception is less vulnerable to the adverse consequences of the problem of logical omniscience than other frameworks we investigate. I point out that various pragmatic features of Stalnaker's framework (the epistemic nature of his possible-world framework, the Gricean principles built into the notion of assertion) help to counterbalance and mitigate the scope of the problem of logical omniscience within it.

In the fifth chapter, *User-friendly Descriptive Semantics*, which is also the first chapter of the second part, I present and motivate my own proposal for the descriptive semantics of proper names and the treatment of modal statements in which they figure. Basically, my aim is to preserve the notion of names as rigidly designating expressions while allowing in as little metaphysics as possible. Working with epistemic possible worlds whose domain is in each case co-determined by a particular context enables me to develop a very intuitive plausible notion of individual, to which a name can be said to refer rigidly. The interpretation of modal statements is then driven not by essentialist constraints in the common sense of the term, but by context-derived limitations, which seems to be a more natural approach.

In the sixth chapter, *Foundational Semantics: Names, Indexicality, and Ambiguity*, I develop a view of foundational semantics inspired by the pragmatic observation that in common parlance one can say that a name, e.g., 'John Smith', can refer to numerous individuals. A lot of attention is given to the ontology of names, and the question 'What is a name?' delivers answers which are then used in a discussion whether names should be seen as indexical or ambiguous. I adopt the ambiguity view, and propose a way of using a Stalnakerian possible-world framework to derive intuitively plausible results for some difficult cases.

In the final chapter, *Conclusion*, I emphasise that my approach throughout the *Guide* relies on a careful examination of different kinds of presuppositions implied by various possible-world frameworks and the notion of the individual used in accounts of the behaviour of proper names. In my own account of these issues, I do without any metaphysical assumptions and aim at describing communication in terms accessible to the speaker and hearer.

Part I

Chapter 1

Outlining the Field and Introducing Some of the Players

1.1 What Should a Theory of Reference of Proper Names Do For Us?

In this thesis, we shall deal with proper names – what they are, how they refer and why, and how is it that speakers know how to use them. These are all questions related to a theory of reference of proper names. Why should one choose proper names? Proper names give rise to many of the problems that make a theory of reference interesting – we have, for example, the Pierre puzzle and the Paderewski case,[1] that bring to our attention the complexities of the behaviour of proper names in belief contexts, and the notorious Hesperus and Phosphorus puzzle that draws our attention to the problems surrounding the informativeness of identity statements and the substitutivity of coreferential names. The aim of my thesis is to provide a critical analysis of existing theories of reference of proper names, to learn from their strengths and weaknesses, and, finally, to attempt to construct a better one.

At the outset we should outline the main issues a theory of reference should address. But before we do that, it shall be helpful to

---

[1]Kripke, 1979

place a theory of reference within the larger framework of a semantic
theory and to say a few words about the relation between a semantic
theory and a theory of meaning.

1.2   Semantic Theory and a Theory of Meaning

Some philosophers (such as Crispin Wright and Garreth Evans[2])
use the expressions 'semantic theory' and 'theory of meaning' inter-
changeably, and I will sometimes follow their usage. Strictly speak-
ing, however, there is an important difference between the two.

Lewis characterises a semantic theory as concerned with certain
relations between words and the world, mainly those on which truth
and falsity of sentences depend.[3] Dummett, on the other hand,
sees this kind of semantic theory as a preliminary step leading ul-
timately to a theory of meaning.[4] According to him, "a theory of
meaning... must say what the speaker knows when he understands
an expression of a language."[5] In Dummett's view, a semantic the-
ory is a part of a theory of meaning, and it "is plausible only in
so far as it provides a basis on which a theory of meaning can be
constructed."[6] We can say therefore that the semantic enterprise
consists of two steps: a construction of a semantic theory, and a
construction of a theory of meaning.

In what follows, we shall see that this dichotomy is reflected in
the characterisation of the task of a semantic theory (in the broad
sense), and of a theory of reference of proper names in particular.

---

[2] Wright, 1993, Evans, 1982.

[3] Lewis, 1972, 169.

[4] "The semantic theory seeks to exhibit the manner in which a sentence is
determined as true, when it is true, in accordance with its composition..."
(Dummett, 1973c, 234).

[5] Dummett, 1973c, 270-271.

[6] Dummett, 1973c, 270-271.

1.3 Descriptive Semantics

An important step in building a theory of reference of proper names is an investigation of their semantics in Lewis's narrow sense.[7] Generally speaking, the task of a descriptive semantic theory[8] is to tell us what the semantics of a particular kind of expression is. It does not concern itself with what it is about the *use* of those expressions that makes a particular semantics the correct one.

The main task of descriptive semantics is to ascribe to the expressions of a given natural or formal language their semantic value, where semantic value is that which provides us with an interpretation of an expression. The crucial theoretical question a descriptive semantic theory has to address is what kind of thing a semantic value for a particular kind of term is.

Following Lewis, Stalnaker, Dummett and others, we can view a semantic theory as concerned specifically with the truth conditions of sentences and with the way the parts of sentences determine their truth conditions. The semantic value of an expression is then seen as the contribution an expression makes to the truth conditions of sentences in which it occurs. In deciding whether a sentence is true, one needs to relate that sentence to the world it targets (in the case of non-modal sentences, for example, the target world is the actual world, while in the case of a counterfactual statement, the sentence should be interpreted in a non-actual possible world). By investigating the truth conditions of sentences in which proper names occur, the task of descriptive semantics is connected with the investigation of the relation between words and the world.[9] The task of descriptive semantics is therefore roughly that of a semantic theory in Lewis's sense.

Should we expect a descriptive theory of reference of proper names to say anything substantial about what should be presup-

---

[7]The rough outline of the following overview is inspired by Stalnaker (Stalnaker, 1997, 534-555), the details and reflections are my own.

[8]This is Stalnaker's expression, adopted from Stalnaker, 1997, 535.

[9]The terms 'word-world relation' and 'word-word relation' are used by Ernesto Lepore in his contribution to Guttenplan, 1995.

posed about the speaker, his intentions and beliefs? Should it say
anything about the use of proper names in a natural language as a
practice? Not really. Answering the query "what makes this expres-
sion refer to that object?"[10] is not the task of descriptive semantics.
Nor is it within the scope of the descriptive inquiry to try and jus-
tify the ascription of a particular semantic value to proper names by
pointing out to the reasons speakers may give for using proper names
in a particular way. Within the descriptive semantic enterprise, we
presuppose language as it is, and try to describe it. In doing so, we
rely on speakers' semantic intuitions to tell what the truth-value of
a given sentence is given a particular state of affairs, and what sort
of truth conditions hold for different kinds of sentences. However
inconsistent this sort of data may be, it forms the basic semantic
facts which one has to rely on in this enterprise.[11]

The descriptive semantics of proper names remains silent on
many important issues: what it is about the practice of using proper
names that endows them with the semantic value they have, what
kind of cognitive value is attached to them, and the role of conven-
tions in all that. Why then should we still want to study descriptive
semantics? Well, the role names play in language and, in particular,
the way they influence the truth conditions of sentences in which
they occur, is perhaps the most readily accessible part of their be-
haviour. Moreover, while our semantic intuitions regarding truth
conditions of sentences can to some degree be manipulated, they
are still among the best guides we have in building any part of a
semantic theory. And this is why the descriptive semantics, even
with all of its limitations, is the cornerstone of a theory of reference
of proper names.

---

[10]Searle, 1967, 93.

[11]On the role of semantic intuitions opinions differ, so that for example
Kent Bach (2002a) sees them as of dubious importance, while Francois Re-
canati (1989) claims that a semantic theory should account for the insights thus
gleaned.

1.4 Foundational Semantics

What makes it the case that a language spoken by a particular individual or a community has a particular descriptive semantics?[12] Answering this question is the task of what Stalnaker, calls a 'foundational semantics'. More specifically, the foundational semantics of proper names looks at what it is about the speaker, her situations, goals, beliefs, and behaviour, that makes it the case that a proper name she uses has a particular semantic value. In this way, foundational semantics is closely connected with a study of the communicative behaviour of speakers.

It is therefore an essential part of the study of foundational semantics to look at how a speaker communicates her mental states (especially beliefs) in her utterances, and to analyse the impact this has on the hearer and her mental states. We shall focus on the particular role that proper names play in communication.

When looking at a communicative situation in which a proper name is used, we can assess the changes in that situation that were brought about by various communicative acts in terms of changes in the intentional states, mainly beliefs, of the participants. Foundational semantics focuses on the verbal interaction between speakers and the resulting changes in their intentional states. That is why we can say that it is a study of the word-word relation. And since it involves intentional states of speakers, one of its aims has to be a creation of a theory that makes plausible predictions about the epistemic situation of speakers. A foundational semantics that fails to be epistemologically plausible (by positing, for example, truth-value gaps for sentences that obviously fulfil their role as assertions in communication) does not meet this criterion.

It cannot be taken for granted that proper names form one homogeneous group. It may turn out that the behaviour of for example geographical names or names of fictional entities is different enough from the rest of the group significantly enough to warrant a separate account. If this were the case, it would have repercussions for all

---

[12]Although this could be viewed as a question about historical background, it is not, as we shall see shortly, meant that way.

parts of a semantic theory of proper names.[13]

Proper names figure in various kinds of speech acts, e.g., questions, assertions, exclamations, and dubbings. It cannot be granted a priori that their semantic value is the same in all speech acts. Most theories of reference, however, focus on assertions, and we shall follow suit somewhat blindly until the last chapter, and section 6.3.2 in particular, where we shall defend a view that different uses of proper names can be explained in terms of the standard, referring use.

1.5   Semantics of Modals

Modal statements have a somewhat special status, and the amount of energy that went into their analysis and study in the last decades is a testimony to it. This is because in analysing modal statements involving proper names, we have to rely not only upon our linguistic intuitions, but also on our preconceptions concerning various kinds of necessity, the identity of individuals, and the nature of possibility. A study of modal statements thus involves both the word-word and the word-world relationship, and has features of both the descriptive and the foundational enterprise.

For our purposes, an attractive theory of modal statements involving proper names is one that makes plausible predictions – and is based on reasonable assumptions – about both the descriptive and the foundational semantics of proper names. In our analysis and evaluation of existing theories of modal statements, it shall be crucial to uncover the implicit and the explicit presuppositions regarding both the descriptive and the foundational semantics of proper names.

In making judgements about the truth or falsity of modal statements, several kinds of intuitions come into play. Semantic intuitions are closely connected with linguistic competence. We rely on them when determining whether a statement is well formed (syntactic intuitions), and when grasping the conditions which would make it

---

[13]I will try to show, however, how various kinds of proper names are connected and defend a view that they do, in fact, form a coherent group, in chapter 6.

true. Some judgements about the truth of modal statements seem to be based solely on semantic intuition (e.g., 'A bachelor cannot be married.' or 'A vixen must be female.'). Judgements about other statements (e.g., 'Franklin D. Roosevelt could have been the son of Theodore Roosevelt.'), however, seem to invoke intuitions of yet another kind – modal ones. Modal intuitions are not based solely on our mastery of language.[14] They express our convictions about the way things are in the world, and about how they could be or could have been. We shall be particularly interested in those modal intuitions that pertain to the potentialities of individuals.

Theories of modality, and especially theories that deal with the reference of proper names in counterfactual statements, usually go beyond a mere construction of a framework for expressing modal behaviour (which would be a metaphysically neutral enterprise).[15] Speakers' modal intuitions tend to be malleable, and one does not have to go too far to produce hypothetical situations where speakers become uncertain or their intuitions clash.

When looking at theories of modal statements, we notice that there is a tendency to predict modal behaviour of individuals, and that descriptions of counterfactuals make presuppositions about the metaphysics of entities involved. A theory of modal statements that does not make metaphysical assumptions about individuals and has plausible consequences for the descriptive and foundational semantics of proper names is hard to come by. When judging whether a particular counterfactual statement is true, we often presuppose that some properties are necessary and others are contingent. In doing that, we express our convictions about the way things are in the world, not just in the language. Modal intuitions have a tendency to differ between speakers. That, and their propensity to be

---

[14]There seems to be some connection between linguistic competence and modal intuitions but it is difficult to spell out how far it goes. For debate on this topic see for example Gibbs and Moise, 1997, and Bach, 2002b. I shall not presuppose any connection between semantic and modal intuitions.

[15]For example, when treating modal statements within a possible-world framework, we could insist on treating the framework itself as a theoretical construct and remain silent on the framework's ontological status.

swayed this or that way by various arguments is the reason why all but the most trivial claims are often hotly contested. This holds especially if our semantic intuitions tell us that a sentence could be true given a possible state of affairs, and a theory claims that that state of affairs is impossible (e.g., Kripke's well-known example of a sentence that is supposed to describe an impossible state of affairs: 'Queen Elisabeth II could have been the daughter of Mrs. and Mr. Truman.').[16]

Why should one want to engage in such a troublesome project? Firstly, we can see modal statements as providing a tough testing ground for our descriptive and foundational theories. If they are plausible, they have to do well for modal statements as well. Secondly, it is puzzling that we have rather clear intuitions about some modal statements (e.g., 'John could have been taller than he is.') but not about others (like 'Sherlock Holmes could have been an unsuccessful detective.'). This whole area is buzzing with questions: Can one argue for a distinction between necessary and contingent properties from language? Do individuals have any non-trivial necessary properties? What is the status of necessary a posteriori statements? Yet even those who try to shy away from providing answers have to be very careful. This is an area where semantics and metaphysics are closely intertwined and easily mistaken for one another.[17]

1.6   The Upshot

It is helpful to keep in mind the distinction between a foundational and a descriptive semantics because it can prevent misinterpretations. Different theories address different parts of a semantic theory, and one should always ask what is the question a particular theory is trying to answer. For example Millian semantics, according to which names have denotation but no connotation, seems like a non-starter

---

[16]Kripke, 1980, 111-113.

[17]For example, direct reference is a semantic claim about proper names. Yet we shall argue that in order to get it off the ground, one has to presuppose haecceitism. But is haecceitism a semantic or a metaphysical doctrine?

to a semanticist interested in speaker's competence, because it does not address the concerns of foundational semantics. Searle's aim "to get at how noises identify objects,"[18] on the other hand, has little to do with the descriptive enterprise.

What are the relations between the three parts of a theory of reference of proper names - the descriptive semantics, the foundational semantics, and the semantics of modals? Their relations are basically of two kinds: going from the theoretical to the pragmatic, the three aspects of semantics are ordered in the sequence in which they are presented here. However, because pragmatic considerations impose restrictions on a theory, we shall see shortly that the three aspects of a semantic theory also impose restrictions on each other in a reverse order.

Semantics of modal statements does not seem quite on a par with the foundational and the descriptive enterprise. Is it really separate from the two? Many theorists claim it is not, but disagree about the side to which it belongs. In Kripke's theory, for example, linguistic and modal intuitions are not seen as really separate, and he seems to regard modal semantics as a part of the descriptive semantics. Other theorists[19] try to make their account of counterfactuals free of descriptive or foundational claims. Theories that view semantics of counterfactuals as a part of foundational semantics also have their advocates.[20] What this shows is that modal semantics can help us shed light on the role of the descriptive and the foundational semantics within a theory. In some theories, as we noted, modal semantics imposes restrictions on the descriptive semantics, in others, on the foundational one. In our own account, we shall return to this issue and argue that modal semantics should not be constructed in such a way that it could impose unreasonable restrictions on the foundational semantics.

What kind of restrictions does foundational semantics place on descriptive semantics? The main condition which a foundational semantics imposes on a descriptive theory is this: if a particular

---

[18]Searle, 1967, 83.

[19]I have Lewis in mind here.

[20]We may view Adams's 'world-stories' (Adams, 1979) that way.

descriptive account is correct, then speakers must, at least for the
most part, know what they are saying according to that particular
descriptive theory. This is so because if one did not know what one
says, one could not mean what one says, which would contradict a
basic assumption that speakers generally mean what they say.

There is a potential mismatch between accounting for what makes
communication successful, and explaining what makes utterances
true. After all, people can, in good faith, make others believe things
that are not true. We can view this as a tension between the world-
word and word-word relationship. Different theories, and directions
of research, try to deal with it in different ways. In the philosophy
of mind, it surfaces as a debate between the externalists and the
internalists or it takes the form of trying to relate broad and nar-
row content. Various answers to the Pierre puzzle, the Twin Earth
thought experiment, and the problem of informativeness of neces-
sary identity statements like 'Hesperus is Phosphorus.' can also
usually be seen as trying to account for the communicative function
of language while getting right the truth of the statements under in-
vestigation. The very fact that this tension between the word-word
and the word-world relation keeps re-occurring may be taken as an
indication that the basic problem is a real one.

Going back to the beginning, we could say that the relation be-
tween a foundational and a descriptive semantics mirrors the relation
between a theory of meaning and a semantic theory, in the sense in
which Dummett uses these expressions. Paraphrasing Dummett,[21]
we could say that a descriptive semantic theory seeks to exhibit the
manner in which the truth of a sentence is determined in accordance
with its composition. This is done by associating expressions with
their semantic values. A foundational semantic theory must go be-
yond this. Its task is to describe what a speaker knows when she
understands an expression of a language, and it has to explain how
the speaker's understanding of an expression fixes a relation between
that expression and the semantic value attributed to it

It is the task of a theory of reference of proper names to propose

---

[21]Dummett, 1973c, 234.

a descriptive, modal, and foundational semantics of proper names such that the descriptive semantics meets the desiderata of the foundational enterprise, and the modal semantics can be made to cohere with both.

1.7 Setting the Scene

For most of the 20th century, the theory of reference was dominated by Frege's and Russell's ideas,[22] which led to the development of descriptivism and the cluster theory of reference. These theories were developed principally by Carnap, Strawson, and Searle.[23] According to them, a proper name has a denotation, which is its referent, and a connotation, a *Sinn* or sense, which is the manner in which the referent is presented.[24] The connotation or sense of a proper name provides not only the means by which the referent is identified but also the meaning of the name. Descriptivism, in all of its forms, claims that there is a semantic relation between a name and the properties commonly associated with its referent.

The underlying motivation of descriptivism is an epistemological one. Descriptivists believe that what is epistemically accessible to us are general properties, such as *being tall, being called 'John Smith', being a featherless biped*. According to this view, our concept of an individual is such that if that individual were stripped of all those properties, properties we use to recognise that individual, there would be nothing left. There is no underlying 'substance' under the properties that distinguish one individual from another. In a theory of reference of proper names, this rather powerful epistemological motivation translates into a claim that a proper name refers to its bearer only as long as that bearer exists, and that a bearer can only be said to exist as long as it has enough of the properties that characterise it in the actual world.

---

[22]Especially Frege, 1893, re-print in English 1952 and Russell, 1905.

[23]See for example Carnap, 1947, Strawson, 1959, Strawson, 1950, re-printed in 1998, and Searle, 1967.

[24]We can imagine the sense as a set of general properties commonly associated with the referent.

In the 1970s, the debate between theories of reference was re-energised by the appearance of new, rather radical theories authored by Donnellan, Kripke, Marcus Barcan, Quine, Stalnaker, Kaplan, and Putnam.[25] These new[26] theories were united in rejecting the then prevalent descriptivism, and in proposing a notion of reference that is not mediated by a term's meaning. The contrast between the new and the old (descriptivist) theories is most apparent in accounts of the reference of singular terms, that is proper names, definite descriptions, and indexicals. This is because part of the motivation of the new theories is a rejection of the epistemological motivation of descriptivism, which was, as we have just seen, motivated by a particular view on the identity of individuals.

It is sometimes said that the new theories of reference, e.g., Kripke's theory, deny that proper names have a *sense*, and that they claim that names are mere 'empty tags', a sort of labels on objects. Speakers are then said to refer not because some conceptual determination of the referent is available to them, but because there is a causal chain that connects their use of a name with the referent. It is because of the stress on a causal connection between the occasion on which a name is bestowed and each particular use of the name, that these new theories are also known as *causal* theories of reference.

If this sketch were an accurate characterisation of the causal theories, they would not have been able to give an account of many of the phenomena that the older theories were rather good at, such as the cognitive significance of necessary identity statements, and would quickly have been abandoned. However, the outline of the causal approach as we presented it now involves a misleading omis-

---

[25]See especially Donnellan, 1966, Donnellan, 1970, Kripke, 1977b, Kripke, 1979, Kripke, 1980, Marcus, 1961, Kaplan, 1978, Kaplan, 1979, Putnam, 1975b, Quine, 1961a, Putnam, 1975a, and Stalnaker, 1984. For more on the debate about the origin of the main ideas of the new theory of reference see Smith, 1995.

[26]I suppose I might use the 'new' here in scare quotes because while they did represent a clear departure from the older descriptivist theories, they have been around for some 40 years, and are not, therefore, as new as they once were. Their newness should be taken with a grain of salt.

sion. What the new theories claim is that while we may need descriptions in order to pick out a referent in a given context, these descriptions do not become a part of the meaning, or semantic value, of the proper name. In this sense, the meaning of a proper name is not co-determined by the name's sense. This is why these theories have been also called theories of *direct reference*. The claim the theories make in the field of descriptive semantics is that the semantic value of a proper name is its bearer, which leaves space for various accounts of the foundational semantics of proper names.

1.7.1 Three Senses of 'Sense'

In order to describe the principal differences between the descriptivist and the new theories of reference more clearly, it is useful to distinguish between the senses in which the term 'sense' can be used.[27] Sometimes the term 'meaning' is used as synonymous with 'sense'. When this is the case, the following distinctions apply to 'meaning' as well as 'sense'.

Sense$_1$ is a purely conceptual representation of an object, which a fully competent speaker associates with his use of a term. The sense$_1$ of a term is something a subject is sometimes said to 'grasp'. It includes only general properties (such as *being male*, *being called 'John Smith'*). Individuals (e.g., John Smith) cannot occur as constituents of sense$_1$.

Sense$_2$ denotes the mechanism that determines the reference of a term with respect to a possible world, time, place, and possibly other contextual parameters. Sense$_2$ is a semantic notion.

Sense$_3$ is the information value of a term, that is, the contribution the term makes to the information content of sentences containing it. Sense$_3$ is a cognitive, epistemic notion. It figures in any belief expressed by a sentence containing the term, and is relevant in determining the epistemological status (a priori or a posteriori, trivial or informative) of such sentences.

Sense$_2$ of a term is closely linked to its semantic value. It is the

---

[27]For these helpful distinctions I am indebted to Salmon (Salmon, 1982, 12).

task of descriptive semantics to describe the sense$_2$ of an expression.
Sense$_1$ and sense$_3$ are related – the former being a psychological
notion, and the latter an epistemic one. Most contemporary se-
mantic theories investigate sense$_1$ of expressions only in the most
general outlines, but many theories deal with sense$_3$ of expressions,
and that sometimes leads to predictions about sense$_1$ and its sta-
tus. In this way, even a purely semantic enterprise may turn out to
make predictions about the psychology of speakers, and these should
be plausible. In other words, while semantics is quite distinct from
psychology, to the extent to which a semantic theory makes any pre-
dictions about the psychology of speakers, these must be plausible.

Sense$_3$ is investigated by the foundational semantics, and an anal-
ysis of modal statements is sometimes employed to highlight the
features of a term's usage that may otherwise remain obscure.

Not all theories of reference of proper names distinguish these
three senses. Descriptivism, for example, conflates all three. It
claims that the descriptions commonly associated with a proper
name (which form its sense$_1$ and sense$_3$) are logically connected (or
even identical) with the meaning of the name (its sense$_2$). Indeed,
this conflation leads to a difficulty in explaining the informativeness
of sentences such as 'Walter Scott is the author of *Waverley*.[28] Some
theorists, Kripke among them,[29] see sense$_2$ as the primary focus of a
theory of reference, while others, for example Stalnaker, take sense$_3$
to be the starting point of their enterprise.

Just like the distinction between descriptive, foundational, and
modal semantics, the distinction between the three senses of 'sense'
shall hopefully prove to be a useful tool for characterising various
theories of semantics of proper names, describing their goals, and
assessing their achievements.

1.7.2   The New Theories of Reference

I mentioned above that what the new theories of reference have in
common is their rejection of descriptivism and a commitment to

---

[28]Searle, 1967, 93.
[29]See Kripke, 1980, 25, footnote 3.

reference unmediated by descriptions. The core motivation of this position is in the conviction that the semantic function of singular terms[30] is to refer to their bearers, to bring their referents, so to speak, into the proposition.

This underlying motivation has been developed in various non-equivalent ways, and brought to bear on different kinds of expressions. This is why it is impossible to identify more features that all of the new theories of reference share, even though there is a sense of kinship between all of them. Kaplan's and Perry's theories[31] use the notion of direct reference and focus on analysing the semantics of indexicals. Kripke developed a theory for proper names and natural kinds, and the key notion of his approach is rigid designation. Having been inspired by Mill, Kripke is sometimes charged with adopting the Millian position with its consequences. We cannot, however, say much more about the new theories of reference without having at our disposal at least a brief outline of the notions that are central to the theories we shall be dealing with.

1.7.3   Introducing Some New Notions

The aim of this section is to give a preliminary characterisation of some notions that play a prominent role in the theories of reference we shall investigate, and show why, regardless of their relatedness, they should be kept apart.

---

[30]At this stage, we can accept Evans' sketch of what singular terms are, given in Evans, 1982, 1-2. Russellian conception of singular terms differs but whenever I would use it in Russellian sense, I would make that explicit. According to Evans's conception, which I shall use, singular terms they include proper names, definite descriptions, demonstratives and some pronouns., What these expressions have in common is their ability to fill in a subject place in a subject-predicate sentence. Other expressions (e.g., 'a man' or 'every boy') can also occupy this position, but they are (normally) not used by the speaker to refer to a particular thing. A proposition containing a singular term is then called a singular proposition. Moreover, according to Kaplan's conception (on which we shall elaborate later, on p. 25 and in section 2.6) singular terms bring the object they refer to, so to say, directly into the proposition.

[31]See for example Kaplan, 1979a, and Perry, 1997.

Direct reference

The notion of direct reference, coined by David Kaplan,[32] has been
used to describe the way demonstratives and proper names refer. It
is also sometimes used as a general name for Kaplan's theory and
theories based on it.[33] These theories vary in the details but share
a commitment to the view that singular terms[34] refer *directly* in the
sense that "the relation between a linguistic expression and the refer-
ent is not mediated by the corresponding propositional component,
the what-is-said, the content."[35] Thus for example, in claiming that
John is tall, one asserts that a property of *being tall* is true of John
– not of anyone called 'John', or a particular individual who bears
that name but of the individual himself. As Kaplan puts it, there is
a sense in which the proposition is about "John himself, right there,
trapped in a proposition."[36] The semantic value of a directly refer-
ential term – the contribution it makes to the propositional content
of the utterance in which it occurs – is therefore just its reference.[37]

In other words, direct reference theories claim that $sense_2$ of a
singular term, its semantic value, is quite distinct from both its
$sense_1$ and its $sense_3$. A direct reference theory makes claims about
the way in which singular terms refer, and its focus is on emphasising
the fact that the semantic function of singular terms is separate
from the means a speaker may use to determine the reference on a
particular occasion.

We should note that an adherent of this view does not have
to repudiate the notion of a meaning ($sense_3$) for singular terms
altogether. Quite the opposite – in explaining the functioning of

---

[32]It is hard to tell where in his writings it appeared first because so much of
his work circulated in the grey circuit for years before being published.

[33]For example Recanati's theory. See his 1993 book called *Direct Reference:
From Language to Thought.*

[34]I use here the general term 'singular term' because while Kaplan's theory
was developed primarily to deal with indexicals, it can, and has been, applied
to the semantics of proper names which, will be my main concern.

[35]Kaplan, 1989b, 568.

[36]Kaplan, 1979, 223.

[37]This is, basically, how Recanati defines the semantic value of a directly
referential expression in Recanati, 1993, 698.

indexicals, direct reference theorists use both the notion of character, which is basically the linguistic meaning of a term, or its $sense_3$, and the notion of content, which we can identify with $sense_2$ of the expression. We shall return to the notion of direct reference and explore it at greater length later, in section 2.5 and several sections that follow.

Millianism

In his 1872 book *System of Logic*, John Stuart Mill proposed a radical view on names. According to him,

> a proper name is but an unmeaning mark which we connect in our minds with the idea of an object, in order that whenever the mark meets our eye or occurs to our thoughts, we may think of that individual object.[38]

The function of proper names is, Mill writes, "to enable individuals to be made the subject of discourse"; names are "attached to the objects themselves, and are not dependent on... any attribute of the object."[39]

This view did not exert much influence for about a hundred years, until Kripke re-introduced it in his 1972 lectures on *Naming and Necessity*.[40] Mill's claim, as presented by Kripke, is that proper names have denotation, but not connotation, and they are but a sort of label on the object they name. Their meaning ($sense_2$) is their reference. This view became known as Millianism – it is not intended to be a faithful rendition of Mill's views. It is rather just the view that Kripke ascribes to Mill.

Millianism is ambiguous between two positions: either names do not have a $sense_1$ and $sense_3$, or else the $sense_1$ and $sense_3$ of names are synonymous with their $sense_2$. We should notice that this view tells us not only what the semantic value of a proper name is – its referent, but also what its linguistic meaning is – if there is any

---

[38]Mill, 1872, reprint 1949, 22.
[39]Mill, 1872, reprint 1949, 20.
[40]Kripke, 1980, 26.

meaning to speak of, it is equivalent to the sense$_2$, i.e., the reference. Even on this brief characterisation it is clear that Millianism presents a far more radical claim than the direct reference view, which is not hostile to meanings (sense$_1$ and/or sense$_3$) being ascribed to proper names. The two views share the focus on the referring relation, but lead to different views on the meaning of names.

Rigid Designation

The notion of rigid designation arises in connection with the modelling of the behaviour of modal statements in terms of possible worlds. At this point, we shall assume that a possible-world framework is neutral with respect to the issues regarding the semantics of proper names, and use it simply as a tool for investigating the truth values of modal statements.

Rigidity is a claim about the behaviour of certain terms, in our case proper names, with respect to possible worlds.[41] Rigid designation is a property a designator (e.g., a name) has if "in every possible world it designates the same object."[42] As a consequence, a referring proper name is said to refer to the referent it has in the actual world necessarily, that is, even in counterfactual situations.

This well-known definition leaves some relevant questions unanswered. First of all, we should ask what happens in the counterfactual worlds where the referent does not exist? There are two possibilities: Either, the designator is assumed to refer to the same object in all the worlds in which the object exists, and otherwise to nothing, in which case following Nathan Salmon, we shall call such a designator *persistently rigid*, or else, the designator designates the same object with respect to all possible worlds, regardless of whether the object exists there or not. Following Nathan Salmon again, we call such a designator *obstinately rigid*.[43]

---

[41] Other kinds of expressions, for example natural kind terms, are also claimed to be rigid designators but we shall focus on proper names.

[42] Kripke, 1980, 48.

[43] Salmon introduces these terms in Salmon, 1982, 33-34. However, the problem was explored even earlier by Kaplan, in Kaplan, 1989b, 569-571 and other

We can visualise the difference between these two kinds of rigid designators using the notion of intension. Given a proper name, we may consider a function that assigns to every possible world the referent of that name with respect to that world. This is the function Carnap[44] calls 'intension'. Using intension, we can distinguish between obstinate and persistent designators by noting that the intension of an obstinate rigid designator is a *total* constant function on the domain of possible worlds, while the intension of a persistent rigid designator is a constant function that is undefined for the worlds in which the referent does not exist. The intension of a persistent designator is therefore a *partial* constant function.

Secondly, we may also ask what happens if a name has no referent in any possible world. In his unpublished *Reference and Existence* (John Locke Lectures of 1973) Kripke introduces a special designator to account for the reference of such names. A *rigid nondesignator*, he says, designates nothing with respect to every possible world.[45]

Finally, there is the special case of entities that are said to exist in every possible world.[46] To account for the reference of these names, Kripke introduces the notion of a *strongly rigid designator*. A designator is strongly rigid iff it designates something necessarily existent, i.e., something that exists in every possible world.[47]

A little overview shall help us keep in mind the different kinds of designators we have introduced:

---

places, who noticed the potential ambiguity in Kripke's definition.

[44]Carnap, 1947.

[45]I return to this topic later when treating names of fictional characters.

[46]This is not really relevant for proper names, since their bearers always exist only contingently.

[47]For Kripke, the names of numbers constitute examples of strongly rigid designators. This is a consequence of his views on the nature of mathematical objects, views that are largely independent of his views on natural language semantics.

Kind of designator	definition	example
Non-rigid designator	may refer to different entities in different possible worlds	'the president of USA'
Rigid designator	refers to the same entity in every possible world	'Julius Ceasar'
Strongly rigid designator	refers to the same entity in every possible world and its referent exists in every possible world	'2'
Rigid nondesignator	refers to nothing with respect to every possible world	'unicorn'

We can further distinguish two different kinds of rigid designators depending on the origin or source of their rigidity. According to this distinction, an expression is rigid *de jure* if it is rigid in virtue of belonging to a class of expressions that are rigid designators, that is, if it is rigid solely in virtue of its (descriptive) semantics. Thus for example proper names are usually seen as rigid *de jure* because their semantics is such that they refer rigidly to their bearers. However, there is another kind of rigid designator that does not refer rigidly in virtue of belonging to a class of expressions which are rigid designators but due to some other kind of fact. These expressions are then called rigid *de facto*. Just which expressions are rigid in this way depends on the line one draws between the semantic facts and other kinds of facts. For example, if we claim that mathematical functions behave the way they do because mathematics deals with necessary facts, then the expression 'the result of adding two and three' will be rigid *de facto* because of our views on the nature of mathematical functions.

In the following table, we have a little overview of the ways in which a designator can be rigid:

Persistently rigid designator	refers to the same entity in every world where the entity exists, otherwise it does not refer
Obstinately rigid designator	refers to the same entity in every possible world regardless whether that entity exists in that world or not
Designator rigid *de jure*	rigid in virtue of its (descriptive) semantics, e.g., 'Julius Caesar'
Designator rigid *de facto*	rigid in virtue of a non-semantic kind of fact, e.g., '3+2'

What we have seen now is that while the notions of rigid designation, direct reference, and Millianism are related, they focus on different parts of the relation between a referring term and its referent. Millianism seems to be a claim about the sense$_1$ and sense$_3$ of proper names. It is denied that proper names have such senses. 'Direct reference' characterises a particular way in which a designator refers. An expression refers directly if its reference is not mediated by the meaning of the expression (its sense$_1$ or sense$_3$). It is a claim about the descriptive semantics of proper names. Rigidity is a claim about a term's referring behaviour across possible worlds, and while it may *prima facie* look like a descriptive claim, we shall see that some of its preconditions have relevance for the foundational semantics of proper names. It is important to note that a designator can be rigid without being direct (for example the aforementioned 'result of adding two and three'). A Kripkean characterisation of the referring relation, which uses only the function of referring in the domain of possible worlds, does not allow us to distinguish between directly referential expressions and those that rigidly refer for some other reason. That is why Kaplan thinks that direct reference is what underlies rigid designation, saying "how could rigid designation not be based on some deeper semantical property like direct reference?"[48]

---

[48]This he says in the following context: "Direct reference was supposed to provide the deep structure for rigid designation, to underlie rigid designation, to explain it... How could rigid designation not be based on some deeper semantical

Epistemic versus Metaphysical Distinctions

It used to be a part of received wisdom that whatever is known a posteriori must be metaphysically contingent, and whatever is known a priori must be metaphysically necessary. This view was based on a partial misunderstanding of Kant's position in his *Critique of Pure Reason*.[49] Kant contributed to this misunderstanding by claiming, for example, that "...if we have a judgement which in being thought is thought as *necessary*, it is an *a priori* judgement," and asserting that "experience teaches us that a thing is so and so, but not that it cannot be otherwise."[50] Regardless of what Kant might have meant, in many schoolbook versions of his views the epistemic and metaphysical modalities were coupled and assimilated into each other.[51]

One of Kripke's most influential undertakings was to revisit the two kinds of distinctions: the metaphysical and the epistemic one, to re-phrase their definitions in new terms, and to separate them clearly.[52] As before, the distinction of a priori versus a posteriori is an epistemological one. Something is knowable a priori if it can be known on the basis of purely a priori evidence. The emphasis is on the *can*: there are facts that can be known a priori to some people, but only a posteriori to others.[53] Something is knowable a posteriori if it can be known on the basis of a posteriori, empirical evidence.[54]

Necessity and contingency are metaphysical terms – if something

---

property like direct reference? It could not be an *accident* that names were rigid and descriptions were not." (Kaplan, 1989b, 570-571).

[49]Kant, 1787, 1963 edition, Kant was not the first to investigate the relation between a prioricity and necessity. Hume, in Hume, 1748, also dealt with this issue but Kant's views on this topic have been more influential.

[50]Both quotations in Kant, 1787, 1963 edition, B3.

[51]My source were my first philosophy textbooks, which were illegal samizdat, no doubt yellowed and crumbled by now, but among the more up-to-date sources, even the entries of the on-line Stanford Encyclopaedia of Philosophy seem to lend themselves to this misinterpretation.

[52]Kripke, 1980, 34-41.

[53]Kripke, 1980, 34.

[54]Kripke, 1980, 35. This characterisation of the terms a priori and a posteriori leaves a lot to be desired. Just to mention one problem, it is nearly circular. We shall return to it later in much more detail, in section 3.2.6.

is necessary, it is so independently of anyone's knowledge of that fact. In Kripke's theory, the two notions are modelled with the help of a possible-world apparatus. Necessity is ascribed to those properties that are true of an entity in every possible world where that entity exists. Properties are contingent if an entity has them in some possible worlds but not in others.[55]

The emphasis on the separation between the metaphysical and the epistemological distinctions gave a new lease of life to some traditional questions, such as whether there might be necessary a posteriori truths or contingent a priori ones. One of Kripke's most important claims is that there indeed are necessary a posteriori properties, and we shall spend a considerable amount of time and space (in section 3.2.8 and several sections that follow) investigating just how he argues for it.

Essentialism

In modern times, necessity and contingency have often been thought to be properties of states of affairs or, derivatively, of statements that express them. Kripke, on the other hand, subscribes to the view that properties too can be contingent (i.e., accidental) or necessary (i.e., essential).[56] He subscribes to essentialism, to "the view that certain properties of things are properties that these things could not fail to have, except by not existing."[57]

Kripke[58] argues that if you refer, for example, to the winner of the 1972 US presidential election, you refer to Nixon contingently because someone else might have won the election. He says that "when you ask whether it is necessary or contingent that *Nixon* won the election, you are asking the intuitive question whether in some counterfactual situation, *this man* would in fact have lost the election."[59]

---

[55]Kripke, 1980, 40-41.

[56]The distinction between accidental and essential properties can be traced back to Aristotle's *Topics*. It was commonly used for most of the Middle Ages, adopted by Thomas Aquinas, Abelard, Albertus Magnus and many others.

[57]As Salmon, 1982, 82 puts it.

[58]Kripke, 1980, 40-42.

[59]Kripke, 1980, 41, emphasis in the original.

Intuitively, Nixon could have lost the election. Winning the election is therefore Nixon's contingent property. He would still be the same individual had he lost. The point is that we can ask about an actual-world individual whether it could, in a counterfactual world, have other properties than it has in the actual world, and we can do this without having necessary and sufficient criteria to identify (or re-identify) that individual in counterfactual situations. We can do this as long as the property we consider is contingent to the individual in question – as the property of *winning the 1972 presidential elections* turned out to be contingent to Nixon. Without going into anything like criteria for being Nixon, we can, as Kripke puts it, "point to the *man*, and ask what might have happened to *him*, had events been different."[60]

This position presupposes a rather innocuous form of essentialism, where nothing more that an essential property of *being Nixon*, further unspecified, needs to be assumed. This is the most moderate form of essentialism that is needed if rigid designation is to function for proper names.[61] However, we shall show that Kripke defends also a rather more robust form of essentialism, in which he proposes the existence of much more specific essential properties (such as *having the parents one actually has*). We shall return to this topic (in section 3.2.9 and following sections), analyse Kripke's reasons for adopting non-trivial essential properties, look into the consequences of his position, and evaluate under what presuppositions it may or may not be defensible.

Haecceitism

The idea that there may be properties which an individual must have if it is to exist, i.e., to be the individual it is, is common to both essentialism and haecceitism. Most authors who adopt essentialism are ultimately interested in describing what non-trivial properties are necessary of an individual or a class of individuals. David Kaplan

---

[60]Kripke, 1980, 46.

[61]We shall investigate the connection between essentialism and rigid designation later in greater detail.

is an exception. He tries, rather successfully, to circumvent the problems that a non-trivial essentialism creates in the foundational semantics of directly referential terms by using a form of trivial essentialism known as *haecceitism*. In this way, while availing himself of some of the systematic advantages of essentialism (like accounting for trans-world individuals), he avoid having to make any claims about particular non-trivial properties that may be essential to an individual or a class of individuals.

Kaplan borrowed the term *haecceity* from Duns Scotus, and just like Millianism is not an effort to reconstruct Mill's doctrine, haecceitism is not a reconstruction of Scotus' views. Haecceitism is introduced as a claim that a common haecceity (which literally means 'thisness') may underlie extreme dissimilarity. A haecceity is, according to Kaplan, that what allows us to

> . . . speak of a thing itself - without reference either explicit,
> implicit, vague or precise to individuating concepts (other
> than being *this* thing), defining qualities, essential attributes,
> or any other of the paraphernalia that enable us to distin-
> guish one thing from another.[62]

Each individual has its particular haecceity, uniquely instantiated in it. It is not to be confused with the property of *being identical with itself*, which is a general property.

Haecceitism is useful in accounting for the reference of a directly referential term in counterfactual situations. Its claim is that

> . . . it does make sense to ask - without reference to common
> attributes and behavior - whether *this* is the same individual
> in another possible world.[63]

Haecceitism thus provides us with the means to say that an individual in one possible world is identical with an individual in another possible world in the literal sense of being one and the same individual, irrespective of properties gained or lost. Haecceity is what

---

[62]Kaplan, 1975, 723.
[63]Kaplan, 1975, 722.

guarantees an individual's identity across possible worlds.[64]

The opposite of haecceitism would be an anti-haecceitist view, a claim that although our interests, such as an intuitive interpretation of counterfactual statements, may lead us to provisionally posit an identity between entities that exist in different possible worlds, these entities are strictly speaking distinct.[65] On an anti-haecceitist view, when we interpret a statement like 'Nixon could have finished his term in office' we intuitively link the actual-world Nixon with the Nixon of the counterfactual world where he was not forced to resign, but the two Nixons are not identical. They share a lot of properties, like *being called 'Nixon'*, *having massive eyebrows*, etc., but the link is in principle the same as the link between all individuals in the actual world who share, for example, the property of *being an American president who did not finish his term*.[66] The anti-haecceitist view repudiates the idea that there may a property that uniquely characterises the individual who possesses it (its haecceity). We seem to be left with a view that individuals are nothing beyond their general properties.

Even on the basis of this short exposition,[67] we can see that haecceitism is a tool for securing a reference for singular referring terms across possible worlds. It is, in fact, a form of an (almost) trivial essentialism, making a commitment only to a further unanal-

---

[64]We should note that haecceitism expresses a view on identity of individuals that is the very opposite of that taken by descriptivists, as we characterised it on p. 19.

[65]Lewis's theory of counterparts, which we introduce in section 3.1.2 is a good example of an anti-haecceitist view.

[66]To the best of my knowledge the property in question - *being an American president who did not complete his term in office* - is shared by (1) Nixon who resigned, (2) the four assassinated presidents: Abraham Lincoln, James Garfield (an amazing gentleman who apparently entertained his friends by writing simultaneously Greek with one hand and Latin with the other), William McKinley, and John Fitzgerald Kennedy, and (3) the presidents who died a natural death while in office: William Harrison, Zachary Taylor, Warren Harding, and Franklin Delano Roosevelt. (We do not count the vice-presidents who finished the presidential term and were not re-elected.)

[67]And we shall return to haecceitism and its opposites in section 3.2.2 and the following section.

ysed property unique to each individual. However, it is still a view on the identity of individuals, and as such it has, as we shall see, consequences for the foundational semantics of directly referring expressions.

1.8 Conclusion

We have now introduced some of the distinctions and terms that will help us in subsequent investigations. We will keep in mind the difference between rigid designation and direct reference, and between essentialism and haecceitism. The division of a semantic theory into a descriptive and a foundational part, with the semantics of modal statements kept provisionally separate, will provide the backbone of the structure of this thesis. We shall fully appreciate the importance of separating a semantic theory into different parts as we shall proceed with our inquiry. The different parts of a semantic theory not only deal with different issues but are intended to answer different questions. The emphasis on looking for the question or a problem that motivates a particular part of a theory shall be an important part of our method.

We have also introduced some of the notions which we shall analyse later in much more detail. However preliminary these characterisations are, they give us a certain platform on which we can work. We shall use them now to characterise the theories from which they originated.

Chapter 2
Descriptive Semantics

2.1 Introduction

The aim of this chapter is to present the most important claims
made by the new theories of reference in the field of the descrip-
tive semantics of proper names. We shall proceed in a systematic
fashion, taking one claim after another, looking at the arguments
proposed in its favour, and evaluating them. Most of the claims
examined here originate in Kripke's and Kaplan's work. Needless
to say, the picture of reference present here has exerted large in-
fluence on the whole environment of the philosophy of language, as
witnessed by the work of Garreth Evans, Gilbert Harman, Leonard
Linsky, John Perry, Hilary Putnam, Francois Recanati, and Gabriel
Segal, to name just a few.[1] Many authors have provided excellent
criticism of the new theory of reference, and their work stands ex-
plicitly (like Salmon's and Stanley's work[2]) or implicitly (e.g., Jay
Rosenberg's and Francois Recanati's work[3]) in the background of
this chapter.

We shall address a number of questions connected with the no-

---

[1]See Evans, 1973, Harman, 1986, Linsky, 1977, Perry, 1980, Putnam, 1981,
Recanati, 1989, Segal and Larson, 1995.

[2]Salmon, 1982, Stanley, 1997.

[3]Rosenberg, 1994, Recanati, 1989.

tions presented above and their role in the descriptive semantics of
proper names: What is needed to derive the familiar Kripkean pic-
ture of names? Why should we think that Fregeans are wrong about
proper names? Why should we suppose that names are rigid des-
ignators? Why should we assume that names refer directly? What
are the ontological assumptions and consequences of direct reference
and of rigid designation? What is the relation between the linguistic
meaning and the content of a directly referring expression? What
are the immediate consequences of these descriptive claims for the
modal and foundational enterprise?

2.2   Arguing for Rigid Designation

In what follows, I review some influential arguments that have been
proposed in favour of rigid designation. In the literature,[4] we some-
times find the arguments for rigid designation classified into three
groups: the modal, the epistemological, and the semantic ones. I
shall follow this division. For the presentation of these arguments
I rely extensively on Nathan Salmon's work.[5] We should keep in
mind that the following arguments apply to proper names as used
in a particular context. This assumption allows us to avoid here the
issues that might be connected with the fact that there is a sense
in which a proper name, such as 'John Smith', can name more than
one individual.[6]

2.2.1   The Modal Argument

The modal argument for rigid designation of proper names is gen-
erally attributed to Kripke. He is interpreted as following this line
of reasoning for example in his Nixon example.[7] Using Salmon's

---

[4]For example in Salmon, 1982, 23-30 and in Stalnaker, 1997, 548-553.

[5]Salmon, 1982, 23-30.

[6]We shall closely examine what happens when one drops this assumption
later, in chapter 6.

[7]Kripke, 1980, 40-46.

presentation,[8] we shall look now at what a fully developed version of this sort of argument looks like:

Consider the name 'Shakespeare' as used to refer to the famous English playwright and poet. Consider then the properties that someone – a literary historian, the author of the *Naming and Necessity* or the shortest spy – might associate with the name and take them to form its sense. These properties could include Shakespeare's distinguishing characteristics such as being the author of *Hamlet*, *Macbeth*, and *Romeo and Juliet*, or being a famous English poet and playwright of the late 16th and early 17th century. Fregeans would admit into this list only general properties, such as those mentioned above. They would thus exclude such properties as Shakespeare's haecceity or his property of *being this very individual*.

Suppose then that the name 'Shakespeare' means 'the person, whoever he might be, who was a famous English poet and playwright of the late 16th and early 17th century and the author of *Hamlet*, *Macbeth*, and *Romeo and Juliet*' or, for simplicity, 'the English playwright who wrote *Hamlet*, *Macbeth*, and *Romeo and Juliet*'. Suppose further that the Fregeans are right and the name is descriptional in terms of these properties. Consider then the following sentences:

> (1) Shakespeare, if he existed, wrote *Hamlet*, *Macbeth*, and *Romeo and Juliet*.

> (2) If anyone is an English playwright and is the sole author of *Hamlet*, *Macbeth*, and *Romeo and Juliet*, then he is Shakespeare.

If the Fregean descriptivism is correct, sentences (1) and (2) express the (descriptive) sense of the name 'Shakespeare' and are therefore true in virtue of their meaning. That implies that they express necessary truths, and are thus true in all possible worlds.

However, Kripke would argue that sentence (1) is not necessary: Shakespeare surely might have lived without ever writing *Hamlet*, *Macbeth*, and *Romeo and Juliet*. He could have become a butcher, a

---

[8]Salmon, 1982, 24-27.

baker or candlestick maker. Sentence (1) therefore does not express
a necessary truth. Sentence (2) does not fare much better: someone
else, for example Francis Bacon, might have written the plays in
question. It is thus certainly not impossible that someone other than
Shakespeare could have been the author of these plays. We can thus
conclude that neither sentence (1) nor sentence (2) is necessarily
true, which they would have to be if they were descriptional in a
Fregean way. Therefore, the name 'Shakespeare' is not descriptional
in the Fregean sense.

According to a modified version of descriptivism, the *cluster the-
ory of reference*[9] it is not a plain conjunction of properties that forms
the sense of a proper name. The sense of a proper name is identified
with a disjunction of (sets of) properties. This disjunction then takes
on the reference-determining function. Sentences (1) and (2) then do
not have to be necessarily true. The modal argument, however, can
be adjusted to take this into account. We can replace sentence (1)
by something like 'Shakespeare, if he existed, either wrote *Hamlet* or
wrote *Romeo and Juliet* or was a distinguished playwright and poet
of the late 16th and early 17th century... Even on the cluster theory,
this sentence is necessarily true in virtue of its meaning. And from
here on, the modal argument can proceed in a more or less similar
way as it did against the Fregean descriptivism – Shakespeare could
have lived without ever having done any of those things, etc. The
modal argument can in fact proceed in all cases where a descriptivist
claims that the sense of a name is a complex predicate formed from
general properties that necessarily determine a given individual.[10]

The claim that sentences (1) and (2) are false with respect to
some possible worlds is motivated by the underlying intuition that
the name 'Shakespeare', as used on a particular occasion, denotes
the same person even in those possible worlds where the individual
lacks all of the distinguishing characteristics which can be used to
identify him in the actual world. In this way, the intuition behind the
modal arguments is intimately connected with a particular position

---

[9]Advocated for example in Searle, 1967.

[10]Kripke argues to that effect in the second lecture of *Naming and Necessity*,
(Kripke, 1980, 71-106).

on the reference of proper names, and with views on the identity of individuals.

One might object that there is a weakness in the modal argument: it shows only that proper names are not descriptional in terms of certain simple kinds of general properties. Faced with the modal line of reasoning, some proponents of descriptivism, e.g., Leonard Linsky,[11] have proposed a move to indexed properties, such as the property of holding the *actual* authorship of some work.[12] Sentences (1) and (2) would then be transformed into something like this:

> (1′) Shakespeare is the actual author of *Hamlet*, *Macbeth*, and *Romeo and Juliet*.

> (2′) If anyone is the actual author of *Hamlet*, *Macbeth*, and *Romeo and Juliet*, then he is Shakespeare.

This seems to do the job. If Shakespeare did, in the actual world, write the plays in question, then both sentence (1′) and sentence (2′) are, in fact, necessarily true.

Kripke would claim that there are two problems with this move: Firstly, the property of actual authorship is not the kind of purely qualitative property a descriptivist theory as he defines it, could accept. Secondly, even if this were a successful move against the modal argument, it would not be effective against the epistemological and the semantic arguments.

2.2.2   The Epistemological Argument

The epistemological argument is a version of the Kripke's Gödel-Schmidt example.[13] Let us consider sentences (1) and (2) again. We know from the modal argument that according to the Fregean theory, these sentences should be true in virtue of their meaning, and

---

[11]Linsky, 1977, 84.

[12]*Being the/an actual P* is a property an individual has exclusively in the world in which the utterance takes place. In this context, we assume that all utterances take place in the same world, and that this is then the actual world.

[13]Kripke, 1980, 83-84.

therefore necessary. Assuming the traditional connection between
necessity and a prioricity, they would then have to be knowable a
priori, that is, solely by reflection on the meaning of the concepts in-
volved. If the name 'Shakespeare' were descriptional in the Fregean
way, it should be inconceivable that Shakespeare lived without writ-
ing any of the plays attributed to him or that another person wrote
those plays instead. It should be as impossible to imagine Shake-
speare becoming a baker instead of a playwright as it is to imagine a
married bachelor. This consequence obtains even if the name 'Shake-
speare' were descriptional in the Linsky way, that is, if we could use
reference to Shakespeare's achievements in the actual world.[14]

However, because we might discover that Shakespeare did none
of the things commonly attributed to him, we can conclude that sen-
tence (1) conveys a posteriori information, that is, information that
can be obtained only by recourse to sensory experience. Sentence (1)
therefore cannot be said to convey a priori information, which was
the descriptivist claim. Moreover, because we could discover that it
was someone else, for example Francis Bacon, that wrote *Hamlet*,
Macbeth, and *Romeo and Juliet*, sentence (2) imparts a posteriori
information as well. With regard to the epistemological argument
we thus reach the same conclusion as in the modal argument, namely
that proper names are not descriptional in the Fregean or the Linsky
way, and it is easy to see how the argument could be adapted to hold
against the cluster theory of descriptions as well.

2.2.3   The Semantic Argument

The semantic argument[15] is often thought to be the most persua-
sive of the three. A version of it has been presented by Kripke,

---

[14]If 'Shakespeare' were descriptional in the Linsky way, then still whoever did
not achieve what is commonly ascribed to the actual Shakespeare could not be
Shakespeare.

[15]While I use here Salmon's material from Salmon, 1982, 29-32, I have con-
siderably changed the presentation of this argument, and for any shortcomings
only I should be held responsible.

Donnellan, and Kaplan.[16] As we shall see shortly, Linsky also un-
wittingly provided us with a very good example (involving Thales)
in the course of his defence of descriptivism. Let us look now at
how a semantic argument for rigidity of proper names is supposed
to function.

Consider a set of properties that might be commonly associated
with the name 'Thales', and which, according to a Fregean the-
ory, provide its sense. Linsky says that the sense of a name like
'Thales' can be determined by a simple description, such as 'the
Eleatic philosopher who held that everything is water'.[17] The name
should denote whoever satisfies the description. What Linsky, when
making his point, did not realise is that Thales was not from Elea
but from Miletus, and that he belonged to the Ionian school, not the
Eleatic one. But let us imagine that there was indeed some obscure
Eleatic philosopher who held that all is water. To which of the two
does the name 'Thales' as used by Linsky refer? Most of us probably
feel that it refers to Thales of Miletus. For if it did not, how could
we say that Linsky made a mistake? Linsky, on the other hand,
seems forced by his views to concede that he refers to someone he
has never even heard of.

This kind of argument differs from the modal and epistemolog-
ical ones in that it is not concerned with the question of reference
in counterfactual situations.[18] The semantic kind of argument for
rigid designation explores the principles of reference of proper names
in the actual world. On any descriptional theory, the referent of a
name is whoever has certain descriptional properties. The descrip-
tive theory predicts that in the example above the name 'Thales' as
used by Linsky refers to the obscure Eleatic philosopher. And that
is where descriptivism fails to capture our intuitions about the way
we use proper names.

We have now familiarised ourselves with the three kinds of ar-
guments for rigid designation of proper names, the modal, the epis-

---

[16]Kripke, 1980, Kaplan, 1973, Donnellan, 1970.

[17]Linsky, 1977, 109.

[18]However, this sort of semantic argument could be adapted to make the same
point about the reference of proper names in counterfactual situations.

temological, and the semantic one. In the following section we shall look more closely at what it is that they actually prove, and how well do they do their job.

2.3   Assessing the Three Arguments

The first thing that catches the eye about the three arguments presented above is that they are not really arguments *for* rigid designation – they all are arguments *against* descriptivism. In order to be arguments for rigid designation, it would have to be shown that descriptivism and rigid designation are the only alternatives and that they are exclusive concerning the semantics of proper names. Such argument, however, has not been put forward. These three kinds of arguments can therefore be seen as arguments for rigid designation only in the sense of arguing against the most obvious opposing theory in the field. Keeping this in mind, let us try to assess the arguments on their own merits.

The form of descriptivism the three arguments attack is a rather simple one:[19] it is assumed that the relevant descriptions must be given in terms of general, non-indexed properties, and that every speaker knows the sense of the name in question.[20]

These assumptions, however, do little justice to the various versions of descriptivism that have been proposed and advocated. First of all, while some authors do indeed subscribe to the thesis that descriptions must be given in terms of purely general, non-indexed properties,[21] theirs is not the most sophisticated version of descriptivism. It has often been noted that identifying descriptions must ultimately contain either indexical elements or reference to partic-

---

[19]It is basically a version which Kripke presents in Kripke, 1980, 71.

[20]While it might seem that the properties associated with Shakespeare in sentences (1) and (2) contain reference to individuals, this is not the case - *writing Hamlet*, *writing Macbeth*, and *writing Romeo and Juliet* are general properties.

[21]The Frege-Russell, based on Russell, 1905, and Frege, 1893, re-print in English 1952 position is often perceived as making this claim. This position is described in more detail in section 5.2.

ulars. This view has been defended by some prominent advocates of descriptivism, for example by Peter Strawson and Michael Dummett.[22] In Kripke's presentation of descriptivism, we find no mention of this view. Moreover, as we saw already, once we allow the use of indexed descriptions – which is what Linsky argues for – we cannot use the modal argument against this version of descriptivism.

Secondly, while it has been held (for example by Russell[23]) that a referent should be determined by the descriptions known to a particular speaker, a more social approach has also been proposed. According to Searle,[24] Strawson,[25] and others, what is relevant is not the descriptions an individual associates with a proper name, but the descriptions associated with it in a language community. On this account, a name as used by a speaker refers to its object in virtue of the speaker's participation in a language community.

This modification does not influence the effectiveness of the modal argument. The epistemological argument, however, may be hard to reformulate if one constructs the sense of a name as socially determined. According to this version of descriptivism, no individual speaker has to know the sense of a name to be able to use it successfully. Given the fact that the epistemological argument relied on the notion of a priori information, and a prioricity seems to be a property of individual knowledge, this kind of argument seems to fare poorly once the social aspect of knowledge of identifying descriptions is incorporated into a descriptivist theory.

The epistemological argument may be problematic even if used against an unsophisticated version of descriptivism. It relies on a much disputed connection between the notions of necessity and a prioricity. The strategy used in the argument is to assume that if a statement is true in virtue of its meaning, then it is necessary, and therefore a priori. Let us grant for the sake of the argument that sentences (1) and (2) are in fact necessary. Let us then assume with Kripke that the notions of necessity and a prioricity can be, or even

---

[22]Strawson, 1967, Dummett, 1973a.
[23]In Russell, 1956.
[24]Searle, 1967.
[25]Strawson, 1959, 151.

should be, seen as distinct. Once we do that, we cannot take the step from necessity to a prioricity. As a matter of fact, Kripke himself admits and defends the existence of necessary a posteriori statements and considers the possibility of contingent a priori ones.[26] In short, the epistemological argument relies on making a connection between the notions of a prioricity and necessity that can, and has been, disputed. The argument cannot be made without this assumption, which makes it less that convincing.[27]

The semantic argument is the strongest of the three kinds of arguments. Even on the indexical or 'socialised' version of descriptivism it remains valid. It expresses very clearly the main difference between the two views on the reference of proper names, contrasting the descriptivist view with the directly referential one. As we mentioned before,[28] the descriptivists have tried to express in a theory of reference of proper names their underlying conviction that if we strip an individual of all its distinguishing properties, there is nothing left. Causal theorists have rejected this epistemological motivation, and, just like the descriptivists, found a theory of reference that expresses their position, a view that an individual is something over and above its general properties. One of the core differences between the descriptivists and causal theorists lies thus in the underlying convictions concerning the identity of individuals. The semantic argument does not actually argue for a conception of an individual that makes its identity independent of all characteristic properties. Rather, it sets the conception of individual identity it presupposes against that which is assumed by the descriptivists.

We have seen that the modal and the epistemological arguments

---

[26]For discussions of necessary a posteriori statements see Kripke, 1980, 38, 54-56, 101-10, 112-115, for contingent a priori ones, see Kripke, 1980, 38.

[27]I have been avoiding any mention of analyticity, replacing it by 'being true in virtue of a statement's meaning'. This is because I did not want to get entangled in a discussion of analyticity. While the term has been used in various presentations of the epistemological kind of argument, Kripke is trying to avoid using it, no doubt because he is aware of the disfavour into which the analytic/synthetic distinction has fallen ever since the 1953 publication of Quine's 'Two dogmas of empiricism' (Quine, 1961b).

[28]On p. 19.

are in trouble once the enriched, or more sophisticated, versions of descriptivism are brought in. The epistemological argument is not very effective even against a more basic notion of descriptivism because it employs some problematic assumptions. The semantic argument is the most interesting of the three because it highlights our basic intuitions about the identity of individuals.

2.4 A Direct Argument for Rigidity

We have already noted that the three kinds of argument we had presented argue in fact just against a particular version of descriptivism. In this section, we shall reconstruct a direct argument for rigid designation and evaluate it. The argument we shall present here was proposed by Jason Stanley.[29]

Let us return to the basics of rigid designation. In Kripke's letter to Kaplan, it is defined as follows:

> A designator D of an object x is rigid, if it designates x with respect to all possible worlds where x exists, and never designates an object other than x with respect to any possible world.[30]

This definition has the advantage of being neutral on the issues of obstinacy or persistence of rigid designators.[31]

One of Kripke's most important claims is that in natural language, proper names are rigid designators.[32] We can formulate this claim as

---

[29]Stanley, 1997, 565-569.

[30]Quoted in Kaplan, 1989b, 569, notation changed for consistency.

[31]The problems of persistence and obstinacy of proper names, especially when time is taken into account, are interesting but irrelevant to the issues I address at this point.

[32]We shall use 'D' to stand for a rigid designator, and 'N' for a proper name. The adoption of quotation marks around N is not supposed to suggest that we quote the letter of the alphabet. It should rather serve to remind us that we mention rather than use the names for which it stands.

(**RN**): If N designates x, then N designates x rigidly, where
'N' is replaceable by names of proper names.

The goal of the direct argument for rigidity is to argue for (**RN**). We shall assume that individual variables under an assignment are rigid designators, and that proper names designate at most one thing in a given world.

If we carefully examine Kripke's definition of rigid designation, we can see that there are two ways in which a designator D, which designates an object x, could fail to be rigid:

(a) There is a world where x exists but is not designated by D.

(b) There is a world where x does not exist and D designates something other than x.

If this argument is to succeed, it has to be shown that both of these (a) and (b) are ruled out if D is a proper name, N. Before we proceed, we should note that no separate proof is needed for a case where in a given world, x exists but N designates something else. Given our assumption that proper names designate at most one thing in a possible world, any such situation will be a situation where N does not designate x, and such situations fall under the case (a).

Let us first argue that

(1) if N is a proper name designating x, then in any world in which x exists, x is designated by N.

Suppose that (1) is false, that is, that N designates x, and (a) is true. This is a situation where there exists such an x that it is designated by N and it is possible that x is not designated by N. If this were the case, then an instance of this schema would have to express a possibility,

$$\exists x(x = N \land \Diamond(e!(x) \land N \neq x))$$

'$e!(x)$' means 'x exists', and can be analysed as $\exists y(x = y)$. This means that a case like the following would have to be possible:

> (2) There is someone who is Aristotle but he could exist
> without being identical with Aristotle.

This is intuitively false, and therefore the supposition that (1) is false cannot be maintained. Thus it seems that if N is a proper name designating x, then if x exists in a world, then N designates it.

Let us turn now our attention to claim (b), that is, the assertion that if N is a proper name designating x, then in any world in which x does not exist, N does not designate something other than x. Suppose (b) is true. It would then have to be the case that N designates x and there is a world where the name refers to something else. Then an instance of this schema would have to express a possibility:

$\exists x(x = N \land \diamond(\neg e!(x) \land \exists y(N = y))$

and consequently, something like this would have to express a possibility as well:

$\exists x(x = N \land \diamond(e!N \land x \neq N))$

> (3) There is someone who is Aristotle but Aristotle could
> exist without being him.

Like (2), (3) intuitively seems false. Thus it seems that if N is a proper name designating x, then if x does not exist in a world, N does not designate anything else. We can therefore conclude that the rigid designation thesis concerning proper names is valid.

We should note that if we read (2) and (3) as expressing the contingency of Aristotle being called 'Aristotle', they would be true. Even if this were a possible reading, it would not be the intended one. The argument concerns the behaviour of the proper name 'Aristotle' as it is used in the actual world, that is, as used to refer to Aristotle.

We came to the conclusion that both of the claims (a) and (b) are false. But what does that mean? Once we fix the reference of a proper name in the actual world (that is, once we establish that in our actual use, N refers to an individual x), then in every possible world in which the referent exists it is designated by N, and the name N does not designate anything else in any possible world.

This claim is intuitive and borders on trivial if we take it for granted that we know what it means for an individual to exist in a possible world. That, however, is not a trivial assumption, and we shall see[33] that the ways in which a proponent of rigid designation deals with this issue is not uncontroversial. There is nothing in the definition of rigid designation as we have it now that tells us what it takes for Aristotle to exist in a non-actual world, nothing that would say that Aristotle, who is according to his own definition of *being human* a featherless biped, could not turn out to be, in a counterfactual world a feathered triped.

In the argument, as it is presented by Jason Stanley[34], we may find the apparent identification of the referent, x, with the designator, N, objectionable. It is misleadingly lending the argument more persuasive power than it has. If we translated the second scheme, for example, making the following scenario plausible,

> (3′) There is someone who is called 'Aristotle' but it is possible that someone else could have been called 'Aristotle'.

then the intuitive thrust of the argument would be as good as gone. However misleading Stanley's formulations may be, they are not incorrect, and the argument should be accepted as valid.

However, the definition of rigid designation used here does not address the issue of determining the referent in the actual world. This is not too surprising because that is a topic belonging to the foundational semantics of proper names, while the rigid designation of proper names is a claim about their descriptive semantics. What is surprising is that we can observe that as long as the notion of existence of an individual in a possible world remains unanalysed – as it is here – rigid designation does not conflict with descriptivism as long as we assume Linsky's, enriched version of it. As we saw in section 2.3, this was the conclusion of our investigation of the modal, the epistemological, and the semantic argument, and it is the conclusion we arrive at here as well.

---

[33]Mainly in the course of section 3.2.

[34]Stanley, 1997, 565-569.

Our next task is to look at what, besides the rigid designation claim, is needed to derive the familiar Kripkean picture of rigidly designating and nondescriptive proper names.

2.5 Kaplan's Distinctions

The concepts that we still need in order to derive the Kripkean picture of proper names were introduced by David Kaplan. He realised that the traditional possible-world apparatus does not enable one to fully capture the key insights of the new theories of reference. In particular, he noted that constructing propositions just in terms of truth values in possible worlds is inadequate. In the following sections, we shall get acquainted with the apparatus developed by Kaplan, and see how it helps to solve some of the open questions of the rigid designation picture. At the end of this chapter, in section 2.7, we shall look at Kaplan's own proposal concerning proper names.

2.5.1 Contexts of Use and Circumstances of Evaluation

We have seen that the wording of the definition of rigid designation, in particular the phrase 'designates the same object in all possible worlds' is open to misinterpretation. It is not intended to mean that the proper name in question could not have been used to designate something else. It means that in order to determine the truth value of a sentence containing a particular name in a non-actual world, we have to determine the semantic value the name has in the actual world.

Kaplan generalises this observation, and modifies it so as to account for the behaviour of indexical expressions as well. To this purpose, he introduces the distinction between *contexts of use* and *circumstances of evaluation*.[35] A *context of use* co-determines the content of what is said on a given occasion. For example, when I say

---

[35]Kaplan, 1989a, 494.

'I am at home now', the context of use determines that this sentence expresses the proposition 'Anna Pilatova is at Johan Huizingalaan 154, etc., as of 5:06 p.m., November 15, 2004.' The *circumstances of evaluation* are the particular situation in which we evaluate what was said on a given occasion. The sentence 'I am at home now' will express a different proposition tomorrow evening (if I go through with my plan of going to the movies), and that proposition will, so I hope, be false in the actual world, while true in some counterfactual ones.

Using this distinction, we can say that a rigid designator designates the same object in all circumstances of evaluation given a particular context of use. For example, when we evaluate the sentence 'John could have been called 'Seán'', as spoken in a context where 'John' refers to a particular individual, the proposition expressed by this sentence is true if there is a possible world where the John referred to in the context above is called 'Seán'.

2.5.2   Content and Character

We have now seen that in order to determine the content of a rigidly designating expression, we need first to consider the context of use, and then evaluate the sentence in the particular circumstances that interest us. Just how this works becomes much more clear once we take into consideration that Kaplan focuses not so much on proper names, but on singular terms in general,[36] and on the functioning of indexicals in particular. It will come then as little surprise that the framework he proposes is most easily understood as applied to indexicals.

Given an indexical expression in a particular context of use, the first thing one wants to do is to find out what the content of that expression in that particular context is. In order to describe how this is done, Kaplan introduces a distinction between an expression's content and its character.

---

[36]We have given a preliminary characterisation of singular terms in a footnote on p. 23.

In general, content can be equated with what is said, or the propositional content at a particular occasion.[37] In his early work, Kaplan seems to have defended the view that the content of a singular term is just the designated object.[38] That is, the content of a well-formed singular term D, where D is a demonstrative, an indexical or a proper name, is the appropriate extension of D. In his later work, Kaplan's position seems to imply that the content of a directly referring expression[39] is a function from the circumstances of evaluation to the appropriate extension. This is the function Carnap called 'intension'.[40] Expressions that have a stable, fixed content over circumstances of evaluation are rigidly designating, and their intension is therefore a constant function.

What enables us to determine the content of a singular term in a particular context is the character of that term. The relation between a character and a content of an expression is somewhat like the relation between a term's sense and its denotation in a Fregean picture, in that the character is a way of presenting the content.

> The character of an expression is set by linguistic conventions and, in turn, determines the content of the expression in every context.[41]

In the case of indexical and demonstrative expression, their character is rather easy to establish. It takes a form of rules the knowledge of which can reasonably ascribed to competent speakers. An example of such a rule is "'I' refers to the speaker or the writer', where the phrase 'the speaker or the writer' is not supposed to be a complete description – it determines the agent in that particular context, and that agent forms then the content of the particular occurrence of the

---

[37]Kaplan, 1989b, 568.

[38]There are some variations on this point even within the 'Demonstratives', Kaplan, 1989a, and the 'Afterthoughts', Kaplan, 1989b, and it is difficult to make a systematic sense of it.

[39]We have familiarised ourselves with the notion of direct reference in section 1.7.3.

[40]Carnap, 1947.

[41]Kaplan, 1989a, 505.

expression 'I'. Using a possible-world framework, we can represent characters as functions from possible contexts to contents.

The introduction of the distinctions between contexts of use and circumstances of evaluation and between character and content will help us, as we shall see, to make the concept of rigid designation more explicit and widely applicable. With the introduction of the character/content distinction, we got the means to describe all singular terms as rigidly designating. In the next section we shall see how the use of the distinctions here introduced, together with the notion of direct reference, will help us derive the familiar Kripkean picture of the reference of proper names.

2.6   Singular Terms and Singular Propositions

It is Kaplan's claim that singular terms refer directly that is of most importance to us here.[42] Kaplan claims that regardless of the means we use to determine the referent of a singular term, those means (demonstrations, descriptions, character rules) do not become a part of the content.[43] The content is non-descriptive. This claim directly opposes descriptivism about content. In that sense, it is a stronger claim than rigid designation regarding singular terms, because, as we saw on p. 50, rigid designation is compatible with descriptivism about content.

For the descriptive semantics of proper names, direct reference implies that an object is not *de facto* the referent of a name, that is, its referent in virtue of meeting whatever descriptive conditions we might have used to single it out, but *de jure*, in virtue of the semantics of the name. An example should help us illustrate the difference between rigid designation and direct reference. Suppose we regard the name 'Aristotle' as descriptional in the Linsky way. The name is rigid, but in counterfactual situations its referent will be

---

[42]We have introduced it briefly in section 1.7.3 but it is useful to point to the main features of direct reference here again.

[43]Kaplan, 1979.

singled out on the basis of meeting the descriptive criteria associated with the term, which will link it to its referent in the actual world. On the direct reference view, however, the name 'Aristotle' is not just rigid – it refers, across possible worlds, to Aristotle without any regard to the descriptions that were used to single out Aristotle in the actual world.

The principles of direct reference are mirrored in Kaplan's views on propositions. According to Kaplan, singular propositions, i.e., propositions involving singular terms, are constructed in such a way that the referent occurs in them directly, that is, independently of any means we might have used to single it out in the actual world. Once we determine the referent, it becomes part of what is expressed by the sentence, that is, of the proposition itself.[44]

But how do we know that the individual occurring in one proposition, say 'John is tall' is the same one as that which occurs in another proposition, say 'John could be short'? How do we make sure that even if we evaluate just one proposition in the various possible worlds of our domain, we keep on tracking the same individual, the referent of a singular term present in the sentence? Hopefully, it is clear by now that we cannot do it by recalling the properties that helped us to single out the referent in the actual world. Kaplan's suggestion[45] is that individuals are extended in logical space in much the same way as they are extended in space and time. An individual thus extended through the logical space (e.g., across possible worlds) can undergo countless changes, but we can, and that is crucial for Kaplan's point, "speak of a thing itself - without reference either explicit, implicit, vague, or precise to individuating concepts (other than being this thing), defining qualities, essential attributes, or any other paraphernalia that enable us to distinguish one thing from another."[46] The property of *being this thing* which distinguishes one thing from another is what Kaplan calls 'haecceity'. It is its haec-

---

[44]Kaplan adopts the notion of individuals occurring as constituents of proposition from Russell, who uses that phrase for example in Russell, 1905, 55-56, Russell, 1911, 216-221, and Russell, 1956, 242-243.

[45]Kaplan, 1975, 722.

[46]Kaplan, 1975, 723.

ceity that guarantees the identity of an individual across possible worlds.[47]

We have now finally all the means we need to derive the well-known picture of rigid and non-descriptive proper names. Direct reference is an essential ingredient of this picture because it is this notion that establishes incompatibility with descriptivism in any shape, form or manner. But now that we see the picture, we can also start sensing some problems. For example, when evaluating a proposition across possible worlds we cannot, obviously, drag an individual, say John, around - that would be not just inhuman, but plainly impossible. Individuals are not extended in logical space in quite the same way in which they extend through time and space. In tracking John we cannot rely on his actual characteristics – we have just arrived at the conclusion that once the referent is established, its characteristics are not a part of the propositional content. That is why Kaplan needs to postulate a property that John has necessarily, his primitive 'thisness' or haecceity, which distinguishes him from every other individual. Haecceities do their job of guaranteeing an individual's identity across possible worlds quite well, which is not surprising considering that they were introduced for that very purpose.[48] On the other hand, it may prove troublesome to justify the assumption of haecceities from the perspective of the foundational semantics of proper names but that is one of the topics which we shall deal with later.

---

[47]We introduced the notion of haecceity earlier, in section 1.7.3, in a preliminary way.

[48]Within this particular framework, it is not clear how the reference of a singular term would work in a possible world where the referent does not exist. On the one hand, Kaplan seems to think that names are obstinate designators, thus accounting for the reference of proper names in negative existential statements, such as 'Hitler could have never been born'. Yet on the other hand, if haecceities are what guarantees reference in counterfactual situations, and given that the notion of a haecceity of a nonexistent individual makes little sense, reference in negative existential statements is not accounted for. This is not a problem I want to deal with at the moment, but it is worth noting that this problem remains unsolved.

2.7 Kaplan's Problems With Names

We have seen that the distinctions Kaplan introduces are useful in defining the new view on proper names. Let us now have a look at how Kaplan himself proposes to treat proper names.[49]

What is the character and the content of a proper name? Assuming, as Kaplan does, that names are rigid designators, we can infer that the content of a proper name remains constant over possible worlds. The semantic value of proper names is therefore context-independent – once the actual referent is established, the name has that same referent in every possible world. Regarding the character of proper names, Kaplan says that names "have no meaning other than their referent."[50] Given that the content of a proper name is constant, we come to the conclusion that the character, which is determined by the referring function alone, must then be constant as well. Proper names, therefore, have both a stable content and a stable character.

The character of proper names is given purely by their referring function - names refer directly and rigidly to their bearers. Co-referential names therefore are assigned the same character.[51] The problem with this conclusion is that Kaplan describes character saying that

> ... it is natural to think of [character] as meaning in the sense of that what is known by every competent language user.[52]

Knowledge of character is clearly seen as a matter of linguistic competence. Therefore, if two names are co-referent, a competent speaker should know that this is so. But that is clearly not the case. A speaker who is both rational and competent can claim she has read books by Mark Twain while denying that she every read

---

[49]The view I present here is based on Kaplan's work up to the controversial article 'Words', Kaplan, 1990. I discuss his proposal from 'Words' later, in chapter 6.

[50]Kaplan, 1989a, 562.

[51]Kaplan, 1989b, 598.

[52]Kaplan, 1989a, 505.

anything by Samuel Clemens, and we would want to say that what
she is lacking is knowledge, not linguistic acumen.

In Kaplan's proposal, neither the linguistic difference between co-
referential names, which are distinct words, or the difference in their
cognitive value, such as may come up in belief contexts, is accounted
for. Co-referential names are treated as synonymous, and this is a
direct consequence of proper names having both a stable content
and a stable character. While many theories experience difficulties
in their treatment of necessary identity statements, for Kaplan's
theory this is especially unpleasant because it is quite clearly the
notion of character that fails here. Kaplan is aware of his problem
with proper names. In fact, he goes as far as to admit that

> The problem is that proper names do not seem to fit into the
> whole semantical and epistemological scheme as I developed
> it. I claimed that a competent speaker knows the character
> of words. This suggests that if two proper names have the
> same character, the competent speaker knows that. But he
> doesn't.[53]

The problems described here eventually inspired Kaplan to propose
a very different framework for the treatment of proper names. But
that we shall deal with much later, in the last chapter.

2.8   Conclusion

In this chapter, we tried to establish what is needed to fully deter-
mine a new picture of proper names, according to which names are
rigidly referring and non-descriptive. We examined both direct and
indirect arguments for rigid designation, and concluded that sur-
prisingly these arguments do not rule out sophisticated versions of
descriptivism. The semantic argument for rigidity turned out to be
incompatible with descriptivism but seemed to rely on an implicit
prior assumption about the identity of individuals (in particular,
about what it takes for an individual to exist in a counterfactual

---

[53]Kaplan, 1989a, 562-563.

world). The assumption was that it makes good sense to think of an individual independently of its properties, and that the individuals thus constituted are the referents of proper names. The semantic argument was helpful is contrasting this view with the descriptivist position, according to which it does not make a good sense to think of individuals independently of their properties because there is nothing more to an individual than its properties.

Given that rigid designation *per se* does not exclude descriptivism, we brought in the notion of direct reference, which turned out to be the missing element needed to fully describe the new paradigm of proper names. However, in order to make the direct reference work for proper names, a primitive non-descriptive criterion of identity had to be adopted. This is where haecceitism came in. The function of haecceities was to ensure that when we track an individual through logical space, we keep on tracking the same individual. Kaplan is the main champion of haecceitism, but we shall see that Kripke will have to introduce essentialism into his framework to do the job that haecceitism does in Kaplan's theory. We shall examine that in the next chapter.

Finally, we gave a brief sketch of Kaplan's position on proper names, and came to the conclusion that while his theory has an attractive proposal for the treatment of indexicals and demonstratives, it seems to fail as a treatment of proper names. But Kaplan is aware of it and has proposed a very different framework in his later work, and we shall use that proposal in the last chapter of this thesis.

This chapter was still somewhat introductory. Its goal was to present the descriptive semantics which the new theories of reference propose for proper names, examine some of the arguments that were presented in defence of that view, look a little more closely at the connection between rigid designation and direct reference, and point to the places where presuppositions about the notion of an individual become important. In the next chapter, which shall deal with the semantics of modal statements, we shall go into much more detail in our investigation of the presuppositions of rigid designation, the use of possible worlds for modelling the behaviour of proper names, and identity of individuals.

Chapter 3
Modal Statements, Individuals, and Essences

In the preceding chapter, we saw that in order to arrive at the 'new' view on names, we had to supplement the notion of rigid designation with the assumption of direct reference. Only then is the picture of names as rigid and non-descriptive fully described. It also became apparent that this view on proper names has to be accompanied by assumptions about the identity of individuals. These assumptions are at the core of the dispute between descriptivist theories of proper names and the group of new theories that includes causal theories and theories of direct reference.

An analysis of modal statements, and of the possible-world frameworks that underlie particular approaches, should provide us with a good basis on which to evaluate various positions to the issue of identity of individuals. In this chapter, we shall pay a lot of attention to the transworld identity of individuals but we should keep in mind that the lessons thus learned also have a direct bearing on what is assumed about the identity of individuals within one world, for example the actual one.

I shall present and analyse the possible-world frameworks of David Lewis, Saul Kripke, and Robert Stalnaker, in this order. I chose this particular sequence of presentation because it will allow us to progress from the most realist view on possible worlds to the

least realist one, and – as I will try to show – from one least motivated by speakers' intuitions about natural language to the one that seems to capture them best. We shall focus on the ontological commitments of each of these frameworks, especially those that pertain to the identity of individuals. We shall try to highlight the connections between the way a particular theory builds its possible-world framework, the stance it takes with respect to the issue of transworld identity of individuals, and the predictions it makes concerning the modal status of particular kinds of propositions involving individuals.

In the course of this inquiry, we shall focus on a number of questions, and try to determine the relevant answers with respect to each possible-world framework as we encounter it: What is the author's own perception of the place of modal semantics, that is, what does the author think he models? What presuppositions does his particular position on haecceitism and essentialism entail, and how are those presuppositions argued for? Could this possible-world framework function without essentialist assumptions? How does a particular modal framework co-determine a theory of content, and what is the epistemic status of the terms in which reference is specified? Do Lewis's, Kripke's, and Stalnaker's approaches to possible worlds address the same issues?

3.1 Lewis's Possible-world Framework

In the following sections, we shall examine Lewis's notion of a possible world and his theory of counterparts, and then look at the motivation underlying the two, as well as their mutual relations. We shall pay especially close attention to those parts of his theory that have a bearing on the identity of individuals, i.e., some implications of the theory of counterparts, and the position with respect to essentialism it entails. Finally, we shall assess the plausibility of Lewis's proposal in the context of the broader tasks of a semantic theory of modal statements, focusing on the process of evaluation of modal statements and the role of modal statements in communication.

3.1.1 Lewis: Let's Be Realistic

There are two basic approaches to the notion of possible worlds. One can either be a possibilist, and hold that there are such things as possible worlds, which are entities in their own right, irreducible to anything else, or one can develop a theory that does not require a commitment to the existence of possible worlds. This view is advocated by various actualist and paraphrastic approaches. Theorists who adopt the paraphrastic approach believe that while one may use the terminology of possible worlds, it is, strictly speaking, just a façon de parler, and the apparent reference to possible entities can and should be paraphrased away. Actualists, on the other hand, try to construct possible worlds from actual entities of some kind. Their approach is to try and find some actual entities that are analogous to possible worlds and can, therefore, serve as possibilia.[1]

Our excursion into the ways of building possible worlds starts with possibilism and its most prominent advocate, David Lewis. He summarises his doctrine of possible worlds as adherence to the following theses:[2]

> (1) *Possible worlds exist.* They are just as real as the actual world. They do not actually exist, since to actually exist is to exist in the actual world.

> (2) *Other possible worlds are things of the same kind as the actual world – "I and my surroundings."*[3] They differ not "in kind but only in what goes on in them. Our actual world is only one world among others. We call it alone 'actual' not because it differs in kind from all the rest but because it is the world we inhabit."[4]

> (3) *The indexical analysis of 'actual' is the correct analysis.* "The inhabitants of other worlds may truly call their own

---

[1] This classification is loosely based on Lycan's overview in Lycan, 1979, 285.

[2] My overview is based on Stalnaker's overview in Stalnaker, 1979, 227, which in turn is based on Lewis, 1973, 84-91.

[3] Lewis, 1979a, 184.

[4] Lewis, 1979a, 184.

worlds actual if they mean by 'actual' what we do; for the
meaning we give to 'actual' is such that it refers at any world
i to that world i itself. 'Actual' is an indexical, like 'I' or
'here', or 'now': it depends for its reference on the circum-
stances of utterance, to wit the world where the utterance is
located."[5]

(4) *Possible worlds cannot be reduced to anything more ba-
sic.* Lewis says: "I emphatically do not identify possible
worlds in any way with respectable linguistic entities; I take
them to be respectable entities in their own right. When I
profess realism about possible worlds, I mean to be taken lit-
erally. Possible worlds are what they are and not some other
thing."[6]

This is the basic doctrine, which was introduced in Lewis's 'Pos-
sible Worlds'.[7] How does Lewis motivate this position? The an-
swer to this question comes in two parts: firstly, there is the folk-
psychological motivation which Lewis is overtly trying to offer, and
secondly, there are theoretical considerations that lead him to this
view.

Lewis opens the exposition of his views by saying that "it is
uncontroversially true that things might be otherwise than they are."
Equally uncontroversially, so he continues, we could then say that
"there are many ways things could have been." And concludes that
we could then call them 'possible worlds'.[8]

Lewis is trying to coax us into accepting that we have believed
in possible worlds all along because 'ways things could have been'
are a part of the folk ontology.[9] We should, however, ask ourselves

---

[5]Lewis, 1979a, 184.

[6]Lewis, 1979a, 185.

[7]Lewis, 1979a.

[8]All quotations in this paragraph are from Lewis, 1979a, 182.

[9]It seems clear from the context of the article (Lewis, 1979a) that 'folk on-
tology' is to be understood as being a part of folk psychology. A folk ontology
is the sum of kinds of entities whose existence seems implied by our every-day
ways of speaking and dealing with the world.

whether the alleged existence of something within a folk ontology is a good enough reason for adopting it into a semantic theory. After all, the existence of 'sakes' – as in 'for God's sake', 'for my sake', etc. – is implied by the folk ontology too, and yet we would consider a claim that 'sakes' exist quite eccentric. While folk psychology is a reasonably good guide in our everyday reasoning, folk ontology is not necessarily a good basis for a semantic theory. Even so, let us concede for the argument's sake that we do believe in the existence of the 'ways things might have been'. But still, nothing in folk ontology implies that we should think of them as existing in the same way in which the actual world exists.[10] Folk ontology seems silent on this issue, and Lewis still owes us an argument why we should adopt this view.

Perhaps then thesis (1) is motivated not by an observation of natural language, but by a different kind of consideration: a conviction that modal notions should be analysed in terms of possible worlds because every other explanation turns out to be circular in the end. Let us assume for the time being that 'possibly S' means that S is a consistent sentence. But what is consistency? Rephrasing Lewis,[11] we could say that a consistent sentence is one that could be true (or, equivalently, one that is not necessarily false), the explanation is circular. If a consistent sentence is one that comes out true under some assignment of extensions to its non-logical vocabulary, the explanation is incorrect, because some assignments of extensions are impossible. For example, an assignment that would have the extensions of 'sheep'and 'pig' overlap, is not possible. If, however, we say that a sentence is consistent if true under some possible assignment of extensions, the explanation is once again circular. Lewis argues that if we analyse modal notions as quantifiers over possible worlds but assume that possible worlds are some kind of 'respectable' linguistic entities, for example maximal consistent sets of sentences of some language, or maximal sets of atomic sentences, the theory turns out to be either circular or incorrect, depending on how we explain

---

[10]For example Stalnaker, 1979 defends the indexical analysis of actuality while rejecting Lewis's full-blown realism about possible worlds.

[11]Lewis, 1979a, 183.

consistency.[12] That is why he concludes that possible worlds are entities in their own right, irreducible to anything more primitive. And this, I believe, is the primary consideration that motivates Lewis's thesis (1).

Lewis establishes his ontology in thesis (1), where he claims that possible worlds are just as real as our world, and in thesis (4), which follows from (1), where he states that they are entities in their own right. While thesis (3) does not deal with ontology explicitly, it has an impact on it. It establishes a distinction between actual existence and non-actual existence.[13] As we shall see later, Lewis's apparatus has largely been directed towards an analysis of counterfactuals. This, I believe, can help us understand the very realist approach to possible worlds: If we think of possible worlds rather as if they were counterfactual situations, and view counterfactual situations as situations which would have been actual, had things developed otherwise, then thinking of possible situations as quite on a par with actual situations seems not quite counterintuitive. And taking from there the step into thinking about possible worlds as being of the same kind as the actual worlds is then a natural consequence of this line of thinking.

Now that we have familiarised ourselves with the basics of Lewis's theory of possible worlds, we shall look at how Lewis deals with identity of individuals in possible worlds.

3.1.2   Adopt a Counterpart!

We shall now introduce Lewis's theory of counterparts, and then look at how it relates back to Lewis's notion of possible world. In the preceding section, we saw that once we start thinking of possible worlds as being just as real as the actual world, the conclusion that they cannot be reduced to anything else seems to follow. In this section, we shall try to show that the theory of counterparts is a

---

[12]Lewis, 1979a is largely devoted to debunking less-than-realist ways of constructing possible worlds. The arguments sketched here are found in Lewis, 1979a, 183.

[13]We shall treat this distinction in greater detail later.

natural consequence of the realist approach to possible worlds. At this point, we shall not try to argue for or against realism about possible worlds, but rather just investigate how its various parts are interconnected.

Every possible-world framework takes some view on the issue of identity of individuals. Some theories adopt the notion of transworld individuals, that is, the view that one and the same individual can exist in more than one possible world. Other theories reject that view, claiming that strictly speaking, this cannot be the case. Lewis, who advocates the theory of counterparts, belongs to the latter group.

Lewis outlines the treatment of individuals within his possible-world framework in his theory of counterparts, which he defines by the following theses:[14]

(a) Nothing is in anything except a world.

(b) Nothing is in two worlds.

(c) Whatever is a counterpart is in a world.

(d) Whatever has a counterpart is in a world.

(e) Nothing is a counterpart of anything else in its world.

(f) Anything in a world is a counterpart of itself.

(g) Some world contains all and only actual things.

(h) Something is actual.

The world mentioned in (g) is unique. That is because if something is actual (claim (h)), nothing is in two worlds (b), and we assume the indexical analysis of 'actual', then everything that is in the same world as the actual entity is actual (everything in that world is actual), and from the actual entity's point of view nothing that does

---

[14]Lewis, 1979b, 111.

not share the same world is actual (the world in which the actual entity is then contains all actual things).

Crucial to Lewis's treatment of identity of individuals in various possible worlds is the notion of a *counterpart relation.* Where in a framework opting for transworld individuals I may have different properties in different possible worlds, in Lewis's framework, thesis (b) says that I stay put in one world, my actual world, and have counterparts in non-actual worlds that resemble me in various ways. But my counterparts are not really me.

The counterpart relation is a relation of similarity, and similarity is here understood rather informally. Lewis sees that this may cause problems, admitting that the counterpart relation

> . . . is problematic in the way all relations of similarity are: it is the resultant of similarities and dissimilarities in a multitude of respects weighted by the importances of the various respects and by the degrees of similarities.[15]

Let us, however, put the problems inherent in the notion of similarity aside, and see what else can be said about the counterpart relation. It has a number of important properties.[16] It is nontransitive: If something, x_1, resembles me more closely than anything else in a world w_1, and something else, x_2, resembles x_1 more closely than anything else in its world w_2 does, then x_2 is a counterpart of x_1, but not necessarily my counterpart. There might be something in w_2 that resembles me more closely.

The counterpart relation is nonsymmetric: Suppose there is something, x_1, in a world w_1 that is a blend of my twin sister Marie and me. Suppose also, that it resembles Marie more closely than it resembles me. It may well be the thing that resembles me in that world most closely, and is therefore my counterpart. On the other hand, because it resembles Marie more closely than me, its counterpart in the actual world will be Marie, not I.

Something can have more than one counterpart in another possible world: In a world where there are, for example, two persons

---

[15]Lewis, 1979a, 112.
[16]Which Lewis outlines in Lewis, 1979b, 113.

who resemble me equally closely, both of them are my counterparts. This works also the other way around: there could be a world where Marie and I have one common counterpart, say Marianne, because it holds for both of us that there is nothing in that world that resembles either of us more closely.

As we noted, the counterpart relation is a relation of similarity, and there is little one can say about it in general terms. In particular models, however, we can specify what kind of similarity we are interested in. That is why in a model where the relation is in some way specified, there can be something in a possible world, w_1, that does not have any counterpart in another possible world, w_2, and, by the same principle, there can be some something in some possible world, w_2, that is not a counterpart of anything in a possible world w_1.

Surprising as it is, the counterpart theory is a natural outgrowth of Lewis's conception of possible worlds. We shall try to show this by examining the tenets of Lewis's conception of possible worlds, and focusing on their consequences for the identity of individuals.

We know from thesis (2), (see p. 63) that possible worlds are of the same kind as the actual world, and from thesis (4) that they are not reducible to anything more primitive. This implies, among other things, that if something is a wooden desk in the actual world, some of its counterparts might be wooden desks, and that if I am a person of flesh and blood in this world, then in another possible world my counterpart can still be a person of flesh and blood, and not some kind of shadow of myself (as Quine was once inclined to object).[17] Possible worlds are not some sort of shadows of the actual world; they are concrete to the same degree our world is. (They are also abstract to the same degree, but that is not the point.) The point is

---

[17] "Wyman's overpopulated universe is in many ways unlovely. It offends the aesthetic sense of us who have a taste for desert landscapes, but this is not the worst of it. Wyman's slum of possibles is a breeding ground for disorderly elements. Take, for instance, the possible fat man in that doorway; and again, the possible bald man in that doorway. Are they the same possible man, or two possible men? How do we decide? How many possible men are there in that doorway?" Quine, 1961a, 4.

that if possible worlds have the ontological status Lewis claims they have, they cannot overlap. Here is how we can show it.

Suppose there is an individual, say Anna, in the intersection of the set of individuals of w_1 and that of w_2. Anna is then both in w_1 and in w_2, and by thesis (3) she calls both w_1 and w_2 her actual world. But facts in w_1 and w_2 differ, for otherwise they would be the same world. So something is and is not a fact for Anna, which is absurd. Therefore, an individual cannot belong to the set of individuals of two distinct worlds.[18]

Because possible worlds cannot overlap, they cannot share individuals. If I exist in the actual world, then I cannot at the same time exist in another world because that world would be my actual world. If I am a certain spatiotemporal entity in one world, I cannot be the same entity in another world. The best we can do is to say that in some possible world there is someone a lot like me, my counterpart. And that is what we set out to show – the counterpart theory is a consequence of Lewis's theses (2), (3), and (4) describing possible worlds. The importance of this observation lies in its pointing out that it is not possible to adopt a strongly realist view of possible worlds, and a notion of a transworld individual at the same time.

In the next section, we shall turn our attention to the indexical theory of actuality and some reasoning that seems to make it more plausible than it may seem at first sight.,

3.1.3   The Indexical Theory of Actuality, Natural Remedies, and the Man in the Street

According to the indexical theory of actuality, the actuality of the actual world consists in its being our world, the world in which *this*

---

[18]On a slightly different note, if we discovered that in some far away galaxy there is a world just like ours (except that, perhaps, what seems to be water on that planet is composed of XYZ rather than H_2O), it would still be a discovery about the actual world. Possible worlds are not far-away planets. They are not like Putnam's Twin Earth, which we treat in section 3.3.5.

is being written. Actuality is a property a world possesses not absolutely but relatively, in relation to its inhabitants.[19] This may seem to run counter to our intuition that the actual is, absolutely considered, more real than the merely possible. After all, we care more about actual events than about possible ones, more about actual people rather than possible ones, and more about actual train schedules rather than just possible ones. We aim at exploring and making predictions about the actual world rather that about possible worlds. We seem prejudiced in favour of the actual. This is reflected in our attitudes. When I say 'I intend to meet you tomorrow at 5 p.m.', I am expressing my intention of meeting you in the actual world, not just in any world, for although it certainly is possible that we meet in one or another possible world, what I want is for us to meet in the actual one.

If we do, indeed, think that the actual world is a place fundamentally unlike any other, we may be disinclined to adopt the indexical theory of actuality, and inclined to replace it with one more intuitively plausible. What would such a theory look like? It would probably hold that actuality is a property of the actual world, and that it distinguishes it from all other possible worlds. It is a property the actual world possesses absolutely rather than just in relation to its inhabitants.

This view would take seriously our certainty of our own actuality; it would take into account our prejudice in favour of the actual. It may seem to be a reasonable alternative to the indexical theory. But how does it account for non-actual possible worlds? Presumably, non-actual possible worlds could have been actual, that is, are *possibly* actual. That seems to be the basis of reasoning with counterfactuals. It also implies, however, that for any non-actual possible world w_1, there is some possible world w_2 in which the world w_1 is actual. But then w_1 and w_2 must be one and the same. So w_1 is actual in w_1. In effect, each possible world is actual in itself.

How then does the actual world differ from other possible worlds

---

[19]This much is in the basic tenets of Lewis's doctrine of possible worlds, p. 63.

with respect to the actuality property? It has the actuality property actually, and not just possibly. Yet we just saw that every possible world is actual in itself. So how is the actual world different from other worlds? It is actually actual as opposed to just possibly actual (i.e., actual for its actual inhabitants as opposed to being actual to its possible inhabitants) – but that is the view we wanted to oppose. It thus seems that the 'intuitive' theory of actuality leads to much the same view on actuality as the indexical theory did.

As I mentioned before, an analysis of the functioning of counterfactuals is an important part of Lewis's work,[20] and it can help us understand the reasons behind his adoption of strongly realist possible worlds. It is important to note that when we want to explain our use of counterfactuals, we look not only at the truth conditions of counterfactual statements (e.g., 'This car could have killed me.'), but also at why they express states of affairs that concern us. Counterfactual states of affairs, or possible worlds, concern us because they could have been the case, they could have been actual. Once we take this insight seriously, it seems hard to hold that actuality is an absolute property of the actual world.

It seems that our dissatisfaction with both the 'intuitive' and the indexical theory of actuality stems from an incompatibility of the intuitions which drive our thinking about actuality and possible worlds. We can try to satisfy our 'prejudice in favour of the actual', and run then into problems when explaining why possibilities concern us, or we can start from taking possibilities seriously, and end up dissatisfied with the conclusion that there is nothing inherently special about the actuality of our world. Lewis has a coherent proposal regarding the status of actuality, and unless we come up with a solution that would do justice to both sides of our intuitions, it cannot easily be dismissed.

But even if we consider Lewis's approach to actuality with the seriousness it deserves, we should ask ourselves whether Lewis's 'actual world' is anything like what the proverbial man in the street would think it is. In particular, can we all inhabit the same Lewisian

---

[20]See his book *Counterfactuals*, Lewis, 1973.

actual world? It follows from thesis (3) (p. 63) that I, the writer, inhabit the actual world. I and my surroundings are actual by definition. The same holds for every speaker or writer. But does it follow that my actual world and another speaker's actual world are the same? The answer is not quite straightforward.

One could try to support a positive answer by pointing to the language Lewis uses – he consistently refers to '*our world*', 'the world *we* inhabit'.[21] This, however, is not a very conclusive kind of evidence. We can arrive at a better-grounded answer if we look at Lewis's theory in more detail.

A Lewisian possible world has parts,[22] namely possible individuals. If two things are a part of the same world, we call them 'worldmates'. Being parts of the same world, worldmates are spatiotemporarily related. It works also the other way around – what unifies a world are the spatiotemporal relations of its parts.[23] If things are related in space and time, they are a part of the same world. If something in a world w_1 were to interact with a thing in a world w_2, then the two worlds would have to be identical. This is in part a consequence of thesis (b), introduced on p. 67, which states that nothing can be a part of two worlds. Because distinct possible worlds are not spatiotemporarily related, an event in one world cannot cause an event in another world. There is no transworld causation. And it seems indisputable that other speakers and I are related to one another in space and time, and we do causally interact. Therefore, we inhabit the same actual world. At least in this sense then the world we all inhabit is the actual world of the man in the street.

---

[21]Lewis, 1979a, 184, my italics.

[22]The following presentation follows Lewis, 1986, 69-81.

[23]Lewis here assumes that spatio-temporal continuity defines the identity of a world. It is an independent assumption that cannot be derived from any of the theses that were used to define possible worlds.

3.1.4 Essentialism: A Matter of Choice?

Just as possible worlds are alternatives to entire worlds, so some parts of those worlds are alternative possibilities for individuals. And just as quantification over possible worlds can be restricted by various accessibility relations, so quantification over possible individuals can be restricted by various counterpart relations. We can restrict both of these relations for particular purposes in various ways. We could restrict the possible worlds over which we wish to quantify, and consider only the nomologically or historically accessible, the epistemically accessible, or for example only the doxastically accessible ones. We could impose similar restrictions on the counterpart, relation. We may choose to consider only the worlds that obey the same laws of physics as our world (or the same legal standards, for that matter), and similarly we can consider only counterparts who have the same physiology as we do, or have the same number of toes.

Lewis points out the generality of his framework, saying that

> ... sometimes one is expected to take a position, once and for all, about what is or isn't possible *de re* for an individual. I would suggest instead that the restricting of modalities by accessibility or counterpart relations, like the restricting of quantifiers generally, is a very fluid sort of affair: inconstant, somewhat indeterminate, and subject to instant change in response to contextual pressure. Not anything goes, but a great deal does.[24]

Given this approach, Lewis could accommodate a variety of essentialist positions, including Kripke's. He could restrict the counterpart relation for humans, for example, and consider as a counterpart of a person only something that is human, and that has the ancestry it does in the actual world. He could do the same for material objects and restrict the counterpart relation to only those individuals that have the same material composition. But doing so absolutely, saying that we can always consider only such counterparts, would be an uncongenial move within his framework because Lewis is not an essentialist in the same sense in which for example Kripke is.

---

[24]Lewis, 1986, 8.

Lewis aims at providing a general possible-world framework, more flexible than any of the frameworks we shall consider later. It can be used for all kinds of purposes in response to different kinds of contexts. The counterpart relation is a relation of similarity. Everything is like anything else in some respect, and for particular purposes we choose to consider particular similarities, and, accordingly, adopt various restrictions on the counterpart relation. But it is important to note that no particular restriction on either the kind of possible worlds we wish to consider, or on the counterpart relation, can be motivated from the nature of the framework itself.

3.1.5 A Battle of Individuals: Transworld Versus Worldbound

As we mentioned before, Lewis addresses in his counterpart theory one of the most interesting and difficult questions of modal semantics: Can an individual exist in more than one possible world? The two basic answers on the market are, not surprisingly, yes, an individual can exist across possible worlds, that is, there are transworld individuals; and no, an individual can only exist in one possible world, that is, individuals are worldbound. And we have shown that Lewis is an advocate of worldbound individuals.

In this section, we shall consider one well-known argument against transworld individuals. It does not come from Lewis's writing – for Lewis, worldbound individuals are a natural consequence of his ontology, and he sees little need to provide independent support for his position on individuals. Other theorists, however, have proposed both direct and indirect arguments for and against worldbound individuals.

The argument we shall consider here was proposed by Roderick Chisholm.[25] It is directed against the notion of transworld individuals, and uses reasoning about gradual changes. The principle is simple: Imagine an entity in the actual world, alter its description slightly, adjust the description of other entities in that world

---

[25]I adapted my version from Chisholm, 1979.

to accommodate this alteration, and then ask yourself whether the entity in the world we thus reached is identical with the entity in the world we started in. We start, for example, with the actual William Shakespeare, allow him to live for 53 years instead of 52 years, and accommodate other descriptions in this world, so that Anne Hathaway was married to a man who lived to be 53, etc. Thus we arrive at a description of another world. Let us call the actual world w_1, and the world we arrived at w_2. Is the Shakespeare in w_1 the same man as in w_2? One could object that any identity of the Shakespeare of w_1 with the Shakespeare of w_2 is incompatible with the thesis of indiscernibility of identicals. How could the Shakespeare who lived 53 years be identical with a Shakespeare who lived 52 years?

We might, however, see this as parallel with a different question: How could the Shakespeare who got married at the age of 18 be the same man who wrote *The Winter's Tale*, if the former is young and the latter old? One could perhaps say that it is not true that the old Shakespeare has properties that make him distinct from the young Shakespeare, that it is rather the case that Shakespeare had the property of being young when he got married, and the property of being somewhat advanced in age when he wrote *The Winter's Tale*. His properties, though different, are not incompatible. And this holds for the different possible worlds, too: Shakespeare can consistently have the property of *living for 52 years* in w_1 and *living for 53 years* in w_2. We could thus assume that the actual Shakespeare is identical with the Shakespeare of the world where he lived 53 years.

Now let us now suppose that we arrived at w_2 by not only altering Shakespeare's age, but also introducing alterations to our –actual-world – description of Francis Bacon, Lord Verulam and Viscount St. Albans. In w_2, Sir Francis Bacon lived to be 64 instead of 65, his wife was married to a man who died at the age of 64 etc. Now let us move to a world w_3 where Shakespeare lived to be 54 and Sir Francis 63 years old, while, again, accommodating these changes in the rest of that world. Moving thus from one possible world into the next, we arrive at a world in which Shakespeare died at the age of 65, and Sir Francis at the age of 53. In this world, Shakespeare and Sir Francis

have, so to speak, exchanged their ages. Let us then move into yet other possible worlds and exchange further properties, so that in w_k, Sir Francis authored *Hamlet*, *Romeo and Juliet*, and *King Lear*, as well as the rest of the plays commonly ascribed to Shakespeare in the actual world, and Shakespeare is the Lord Chancellor to James I, the author of *Novum Organum*, etc. Finally, let us move into a world where the two exchange their names and titles. Proceeding in this way, we finally arrive at a world w_n which is like w_1, the actual world, except that the Shakespeare of w_n can be traced back to Sir Francis of w_1 and vice versa. Should we now say that the Shakespeare of w_n is identical with Sir Francis of w_1? In other words, is there such an x so that x is Sir Francis Bacon in w_1 and Shakespeare in w_n? How should we decide?

Assume that we answer in the affirmative: there is such an x such that it is Shakespeare in w_n and Sir Francis in w_1. But if this is the case, how are we to tell the two worlds apart? Should we say that though different, these worlds are indiscernible from one another? The two Shakespeares could be seen as discernible because one has the property of *being Sir Francis* while the other does not. The option of distinguishing worlds by essential properties of some of their individuals shows that there can be a sense of 'indiscernible' on which 'indiscernibles are identical' tells us more than 'identicals are identical'.

If w_1 and w_n are two different possible worlds, then there could be infinitely many other possible worlds as difficult to distinguish from one another as w_1 and w_n are. Why do we assume that the Shakespeare of w_1 is identical with the Sir Francis of w_n? We made this conclusion possible once we conceded that an individual can be found in more than one possible world. It seemed perfectly reasonable to assume that Shakespeare retains his identity through small changes, such as were involved in the transition between w_1 and w_n. These transitions can be as gradual and slight as one likes, but once we allow identity to be transitive, and accept that an individual can exist in more than one possible world, we seem to take the first step on quite a slippery slope.

Is there a way of retaining identity through possible worlds while

blocking the extreme consequences such as we just encountered? Adoption of non-trivial essential properties could be a solution. But if essential properties were to help us with our problem, we would have to be able to specify which properties are essential for a given individual. Only that way could we decide which transitions from one possible world to another lead to a 'loss of identity'. However, Chisholm objects,[26] there seems to be no procedure for deciding which properties are essential and which accidental of a particular individual. Properties can be necessary under a certain description of an object, but, according to Chisholm, we do not have any meaningful procedure that would deliver necessary properties in the required, non-trivial sense, in isolation from a particular context.

We can summarise Chisholm's conclusion as follows: Once we allow for any property of an individual to be changed without loss of identity, there is no acceptable way of stopping the process. Ultimately, we reach the counterintuitive result that, so to say, something can look like an apple, smell like apple, taste like apple, and yet be an avocado. This result could be avoided only by either positing essential properties that are unique to each individual – but, as Chisholm claims, we do not have a good way of deciding which properties should play this role – or by preventing the whole slippery slope altogether. The latter option would amount to banning all changes of properties on pain of loss of identity. This step is in fact equivalent to claiming that all properties are essential of an individual. And if *all* properties are essential to an individual, that includes relational ones as well, and lo and behold, we have just concluded that individuals are worldbound.

The problem with making all properties essential is that modal logic was developed to analyse modal statements, and the predictions a framework of worldbound individuals makes with respect to those statements do not seem to match with our intuitions. Some of us may that think the statement 'If the Twin Towers had not been destroyed in 2001, Bush would not be a president now' is true. But whether this is true is not the point. The point is rather that most of

---

[26]Chisholm, 1979, 85.

us understand this statement as being about the actual George W. Bush, the president of the USA. It is reasonable to see this statement as a claim that had the Towers not been destroyed, things would have gone otherwise for this very individual, George W. Bush (not to mention many other actual-world individuals whose lives would have been different). This reading, however, seems to require George W. Bush to exist in more than one possible world, and that – the possibility of a transworld individual – is what both Chisholm and Lewis argue against.

One may be tempted to point out that if what is needed to stop the argument using gradual changes is a property unique and essential of each individual, we have such properties. They are called 'haecceities'. So why could we not use them? The problem with haecceities is that they could not help decide which properties are essential and which are contingent because haecceities are not descriptive. The argument Chisholm makes could be made even if we assumed that Shakespeare and Sir Francis both retain their respective haecceities.

Chisholm[27] thus seems to prove his point – by allowing the identity of an entity to be preserved through changes, we open the door to some counterintuitive results. Chisholm suggests that in order to prevent these results we should ban all changes in an individual's properties on pain of loss of identity. In the following section we shall see that the problem with his argument might be that it just proves too much.

3.1.6 Can We Survive a Change?

Lewis[28] suggests that we should think of possible worlds as analogical to moments in time. And as a matter of fact, Chisholm's argument is parallel to the well-known puzzle of the ship of Theseus, which deals with the issue of identity of individuals and change over time. We shall now briefly review the Theseus puzzle, and then see which

---

[27]Chisholm, 1979.
[28]Lewis, 1979a, 184.

of its lessons could be applied to Lewis's position and Chisholm's
argument.

This puzzle has been around since antiquity,[29] but for our purposes we shall consider two somewhat modernised versions.[30]

According to the first scenario, Theseus sailed away with a complete supply of new parts as his cargo. While at sea, he gradually replaced each part of his ship with a new one, and threw the old parts overboard. Did he returned to the harbour on the same ship on which he left? Of course it had changed, but was it the same ship?

The second scenario is just like the first, except that Theseus was followed in another boat by a Scavenger who picked up all the pieces as Theseus threw them overboard, and used them to build a boat. The Scavenger then reached the port in a ship composed of the selfsame parts that the ship of Theseus had been composed of when it left the port. He docked his ship next to the ship on which Theseus arrived. Which of the two ships was then the ship of Theseus?

What are our options? Let $\mathcal{A}$ be the ship Theseus started his voyage on, $\mathcal{B}$ the ship he finished it on, and $\mathcal{C}$ the ship the Scavenger finished his trip on. If we were to make a sameness of parts a necessary condition of identity, then $\mathcal{A}$ would be identical with $\mathcal{C}$. That would imply that Theseus changed ships during his voyage, because he started in $\mathcal{A}$ and ended in $\mathcal{B}$. But we know from the story that Theseus stayed on one ship during his whole journey. Alternatively,

---

[29]The story first surfaces in print in Plutarch (Vita Thesei, 22-23) in the following form: "The ship wherein Theseus and the youth of Athens returned had thirty oars, and it was preserved by the Athenians down even to the time of Demetrius Phalereus, for they took away the old planks as they decayed, putting new and stronger timber in their place, insomuch that this ship became a standing example among the philosophers, for the logical question of things that grow; one side holding that the ship remained the same, and the other contending that it was not the same." This reference was found at http://faculty.washington.edu/smcohen/320/theseus.html in October 2003.

[30]The presentation of these two versions was inspired by http://faculty.washington.edu/smcohen/320/theseus.html, obtained in October 2003.

we could claim that since he did not change ships, $\mathcal{A}$ must be identical with $\mathcal{B}$, although they have no parts in common. That would also imply that $\mathcal{C}$ is not identical with $\mathcal{A}$ although it has all of its parts. Both suggestions seem equally counterintuitive.

One might suggest that spatiotemporal continuity is a better criterion of identity than identity of parts. Using this criterion, Theseus arrived at port in the same ship in which he left the port, and $\mathcal{A}$ is therefore identical with $\mathcal{B}$. On the other hand, there are cases when objects are taken apart, and then put together without, at least on the face of it, loosing their identity. If I send a piece of antique furniture for restoration, the piece will be taken apart, exist for a while only as disjoint pieces of wood, and yet I would be very upset if the restorer informed me that the piece of furniture I left with him is gone. I would sue him. Yet, the spatiotemporal continuity that of piece of furniture has been interrupted.

The lesson seems to be that sometimes the criteria we normally use to determine identity break down, and that clear-cut universal criteria are vulnerable to counterexamples. Yet we usually seem to agree that things and individuals can retain their identity through some changes. The problem of identity through change is real, and cannot be brushed aside by claiming that objects do not persist through change.

Lewis is reluctant to accept any form of essentialism. He objects to a distinction between necessary and contingent properties because he thinks it is arbitrary. Yet by adopting worldbound individuals, he makes in effect all properties essential. As we already stated, in Lewis's framework, any change in an individual leads to a change of identity. This framework is extremely general – by adopting various kinds of restrictions on the possible worlds under consideration and on the counterpart relation,, Lewis can model various kinds of modalities (for example, reasoning with logically possible worlds, physically possible worlds, epistemic alternatives). Yet underlying this flexibility, it remains the case that, strictly speaking, individuals are worldbound, and any non-actual scenario, regardless of how small the divergence from the actual world, creates a new possible world and a new possible individual. The Delia Smith who bakes

her apple-pie to perfection and the Delia Smith who burns it are, in this framework, seen as two distinct individuals.

Returning to Chisholm's example: yes, there are real problems that arise in connection with transworld identity of individuals. But these problems are not specific to working in one or another possible-world framework – they are the old problems of persistence through change in a new guise. A move to block the persistence of identity through change is as unacceptable when dealing with transworld identity as it is when dealing with persistence through time. Lewis offers a principled solution to puzzles like that of the ship of Theseus but his solution is too strong. Perhaps the questions of identity and persistence through change are cannot be reduced to black and white. Perhaps they come by their very nature shades of grey.[31]

3.1.7   Looking Into Possible Worlds

Perhaps the problem of transworld identity arises out of a particular way of looking at – or into – possible worlds. Maybe it is a by-product of looking at possible worlds as if they were somehow out there, a view supported to some degree by the way Lewis defines possible worlds and the counterpart relation.,

We saw that on Lewis's view, possible worlds are just as real as our world is, and at least some of them are inhabited by individuals of flesh and blood. At least some of these individuals bear a special relation to individuals in the actual world – they are their counter-parts. But how do we determine what is the counterpart of, for example, my dog Mambo, in a given counterfactual world? Intuitively, we should find in that possible world the individual (or individuals) that most closely resemble the actual dog. Kaplan describes this process[32] as looking through a 'Jules Verne-o-scope.' But suppose that in the possible world we are interested in, dogs have evolved to look just like humans. Which of those human-like individuals is my dog's counterpart? To find out, we take the Jules Verne-o-scope,

---

[31]We shall return to the question of identity in section 5.5 and 5.6 and explain somewhat cryptic remark concerning the colour grey.

[32]Kaplan, 1978, 93.

examine the human-like dogs, and decide which of them is resembles Mambo most closely. As strange as it sounds, on Lewis's view, given that the counterpart relation is a relation of similarity, the problem of locating counterparts is real.

A number of authors, Kripke[33] and Richards[34] among others, argue that locating the counterparts is a pseudoproblem, generated by the 'telescope' view itself, and not by anything in the nature of modality. Kripke points out that possible worlds are our stipulation, rather than something we discover, some far-away lands. Richards raises a related point: How is it possible for us to know anything about those possible worlds given they are 'out there', independently of our mental activity, causally and spatiotemporarily inaccessible to us? He says

> [Lewis's] truth-conditions are such that, for any given modal statement, it is impossible in general to determine whether they are met and hence whether the statement is true. There is, however, a certain measure of agreement between people about the truth-value of certain modal statements. Insofar as there is agreement one must assume that if it is not catechized into the populace without any understanding of any truth-conditions for these modal statements, then there is some other account of truth-conditions for these modal statements, and these truth-conditions are such that we may with some degree of confidence determine whether or not they are met.[35]

If Lewis intends his possible-world framework to give an account of the meaning of modal sentences, and it seems he does, then possible worlds should be such that they can be used in giving an intuitively plausible explanation of why modal sentences function the way they do. In interpreting modal sentences we exhibit a degree of knowledge of what is possible and what is impossible, and that knowledge is

---

[33]Kripke, 1980, 44.

[34]Richards, 1975.

[35]Richards, 1975, 109. We should understand Richards's 'truth-conditions' in a Davidsonian way, where a sentence's truth-conditions form the core of its compositional meaning.

based on our acquaintance with the actual world. Why then should we posit strongly possibilist Lewisian possible worlds at all? Lewis would probably answer that we set out trying to explain our modal notions and our explanans should be independent of our explanandum. Saying that our intuitive understanding of modal statements in any way determines what possible worlds can be, is putting the cart before the horse. This may well be a correct objection, but still, in view of the difficulties Lewis's conception faces, we should perhaps try to find a new equilibrium between explaining modal notions, and giving an intuitively plausible analysis of counterfactuals. Actually, we shall see that extreme possibilism causes problems even in Lewis's analysis of counterfactuals. Let us look at it now.

3.1.8   Counterfactuals and Worldbound Individuals

Lewis provides a well-known analysis of counterfactual statements. It is supposed to be one of the strong aspects of his theory.[36] A statement like 'Had I known how long the tram is going to take, I would have taken the bicycle' is not explicitly modal, but it is not a claim about how things have actually turned out either. In general, statements of the form 'If it were (had been) the case that p, it would be (would have been) the case that q' are about how things could have been, and we can think of them as modal.

Unlike other modal statements, counterfactuals do not require quantification over all possible worlds. In fact, the number of possible worlds that have to be considered is in most cases quite limited. My statement 'Had I known how long the tram is going to take, I would have taken the bicycle', should not be taken to mean that in any world where I know how long the tram takes I take my bicycle. In some worlds, it might be snowing too hard to ride a bicycle. Lewis would say[37] that my claim concerns only those possible worlds that are quite 'close' to the actual world. My statement is supposed to

---

[36]The following argument is adapted from Lewis, 1973, 33-34.
[37]Lewis, 1986, 20-23.

be evaluated in those worlds that are as much like the actual world as is compatible with me knowing how long the tram takes, and it says that in those worlds, I use the bicycle.

On the face of it, this analysis seems attractive. The problem is that our description is not accurate. Strictly speaking, Lewis would analyse the counterfactual statements as being not about me and the bicycle, but as about my counterpart and the bicycle's counterpart., And this does not seem to accommodate our pre-philosophical intuitions about counterfactuals. The statement above could have been followed by an exclamation 'I should have used the bicycle, and that's what I'll do the next time!' Our understanding of counterfactuals can lead to changes in our behaviour, and result in joy or regret. If, however, counterfactual statements are about our counterparts, and not about us, why should we care?[38] As Loux says,[39] "on Lewis's view, things could have indeed gone otherwise, but they could not have been different for any of the individuals existing in our world."

Worldbound Individuals and Proper Names

The worldboundedness of individuals has dramatic consequences for the semantics of proper names. As we just noted, nothing could have been different for actual individuals. If we stay strictly within Lewis's framework, we cannot make good sense of the notion of a rigid designator – it collapses. Lewis does not offer us a useful kind of framework for dealing with issues of reference of proper names. His framework cannot model the tests used to tell for or against theories of proper names: Would Moses still be Moses had he not done anything ascribed to him in the Bible? By Lewis's light, the answer is trivial. It will always be negative. Lewis offers us no guidance here. His framework seems to be of little if any use when it comes to the semantics of proper names. This framework is, as we admitted, very flexible, but its generality, and Lewis's unwillingness to adopt any sort of essentialism, also leads to limitations. The main limitation we have encountered is the worldboundedness of

---

[38]Kripke, 1971 voices a similar concern.
[39]Loux, 1979, 42.

individuals, which results in the inapplicability of rigid designation.

One could, of course, brush this kind of criticism aside as dogmatic, and claim that the counterpart relation is basically a Lewisian equivalent of identity, and that thanks to the generality of his framework, we can restrict it in any way we wish. We could, for example, restrict it so as to model an actual speaker's intuitions about potentialities of individuals as they are reflected in her judgements about the truth and falsity of modal statements. This would stop the counterintuitive examples in which we look with a 'Jules-Verne-o-scope' for my dog's counterparts among humanlike creatures. But even if the counterpart relation was restricted in an intuitively plausible way, the worldboundedness of individuals would be unchanged. It is directly a result of the extremely realist position Lewis adopts. And that seems to be at the core of Lewis's approach to possible world. A change in Lewis's conception of individuals would result in far-reaching changes to the whole framework. We have seen all through the preceding sections that Lewis's theory has an amazingly strong internal coherence. It is impossible to remove the more objectionable parts and preserve the parts that are attractive. Worldbound individuals are not attractive but they cannot be removed without transforming the whole framework beyond recognition.

3.1.9   Conclusion

We have now presented and analysed Lewis's views on possible worlds and possible individuals. In the first two subsections, we reviewed the basic tenets of Lewis's theory of possible worlds and the theory of counterparts., We focused on reconstructing the mutual relations between the theses of each of these theories, as well as Lewis's motivation for adopting them. It is now hopefully clear in what way are both the theory of counterparts and the indexical analysis of actuality a natural outgrowth of Lewis's position on possible worlds.

In particular, we have shown that one can arrive from theses (2), (3) and (4) of the theory of possible worlds at the conclusion that possible worlds cannot overlap or share individuals, and that no

individual can therefore be in more than one world. Then we only needed to add two very intuitive premises (which Lewis lists in the tenets of the theory of counterparts), namely that every individual is in a world, and that something is actual, in order to derive the whole of the theory of counterparts from the theory of possible worlds.

In section 3.1.3 I tried to show why we should not be too quick to dismiss the indexical analysis of actuality. I proposed a hypothetical 'naive' or 'intuitive' theory of actuality, developed it a little to see how it would deal with possible worlds, and concluded that such a theory quickly looses its intuitive plausibility once it is put to work. It may be difficult to propose an analysis of actuality that would work as well as Lewis's indexical analysis does.

Because the counterpart relation is a relation of similarity, not of identity, and everything is like anything else in some respect, there is no reason for Lewis to adopt any sort of essentialism into the general form of his theory. One can restrict the counterpart relation in various ways depending on a particular purpose, but these restrictions are not motivated from within the framework – they are a consequence of the use to which we put the possible-world framework (we can, for example, consider only possible worlds that work in accordance with the laws of physics that hold in the actual world, and restrict the counterpart relation accordingly). Because the counterpart relation is not about identity, Lewis does not need to say anything about criteria of transworld identity.

Finally, we went through a number of objections to Lewis's theory. They all targeted consequences of the counterpart theory. We saw that as it is, Lewis's theory implies a somewhat unusual form of essentialism, according to which all properties are necessary. We also saw that Lewis's framework is not well suited for analysing semantic properties of proper names.

Part of our goal was to demonstrate the internal cohesion of Lewis's framework. This we did, and we arrived at the conclusion that to remove the theory of counterparts, which does seem to have rather unpleasant consequences, we would have to alter the very core of Lewis's theory.

3.2 Kripke's Approach to the Semantics of Modal Statements

In the preceding sections we analysed the problems connected with a strongly realist approach to possible worlds. In the following sections, we focus on a proposal that is markedly less realist about possible worlds, and does not support the notion of worldbound individuals. In this subchapter, we shall reconstruct, as best we can, Kripke's approach to possible worlds, his stance with respect to haecceitism and essentialism, and investigate the role various kinds of essentialism play in his overall plan. Clarifying Kripke's position with respect to these issues should help us evaluate his proposal concerning the semantics of proper names and the metaphysics of modality.

Saul Kripke's first notable achievement was to provide a general framework that treats various systems of modal logic as variations within a common framework.[40] Using the notions of a model, accessibility, validity in a model, and a possible world, Kripke gave a unified semantics to the four main modal systems (system M or K, and systems B, S4 and S5). Where there had been a plethora of seemingly unrelated systems, he brought order and unity.[41] He then went on to apply some of the insights from his work on formal systems to natural language, and it is his work on proper names[42] that we shall focus on right now. A substantial part of his work on the semantics of natural language is focused on the semantics of modal statements that involve individuals. In this analysis, he frequently uses possible worlds; in fact, his famous definitions of necessity, rigidity, and essential property are all given in terms of truth in possible worlds. It is then rather surprising that when we turn to his texts, we find that very little is explicitly said about the notion of a 'possible world' in the context of natural language anal-

---

[40]Kripke, 1963.

[41]He was not the only one. The work of Richard Montague, Stig Kanger, Ruth Barcan, Dana Scott, and Jaako Hintikka has to be mentioned in this context as well.

[42]Kripke, 1971 and Kripke, 1980.

ysis. This may be a consequence of Kripke's formal roots, but in our present inquiry, a formal definition of a possible world leaves many questions open. Filling in the picture of Kripke's notion of possible worlds as it is used in his natural language analysis, that will be our next immediate task.

3.2.1 Stipulating Possible Worlds

It is not difficult to identify some views on possible worlds which Kripke does *not* entertain. We have already briefly mentioned[43] his criticism of what he calls 'the telescope view'. In essence, Kripke criticises Lewis for treating possible worlds as if they were disconnected from the actual world, which he makes clear by pointing out that on a Lewisian approach,

> . . . one thinks, in this picture, of a possible world as if it were like a foreign country. One looks upon it as an observer. Maybe Nixon has moved to the other country and maybe he hasn't, but one is given only qualities. One can observe all his qualities, but, of course, one doesn't observe that someone is Nixon. . . So we had better have a better way of telling in terms of properties when we run into the same thing as we saw before; we had better have a way of telling, when we come across one of these other possible worlds, who was Nixon.[44]

In Kripke's view, the problem results from Lewis's insistence that possible worlds be given by qualitative descriptions. It is the *kind* of description we have of a Lewisian possible world that makes us feel like observers in a foreign country, or – in Kripke's idiom – as if we were 'viewing through a telescope.'[45] On a Lewisian view, we

---

[43]In section 3.1.7.

[44]Kripke, 1980, 43-44.

[45]Kripke, 1980, 44. Actually, the telescope metaphor may have originated with David Kaplan Kaplan (1979c). 'Transworld Heir Lines', though published after Kripke's 1971 Naming and Necessity lectures, was presented as a lecture in 1967. In both the lecture and the paper, Kaplan repeatedly uses the notion of a 'Jules Verne-o-scope'.

know what things look like in a possible world, but not what they
are. This, Kripke goes on to say, not just fails to reflect the way
we intuitively interpret counterfactuals – it breeds unnecessary and
misleading problems concerning the identity of individuals. This was
a view we arrived at in our analysis of Lewis's framework as well.
But let us look now at what Kripke's response to the problem is.

According to Kripke, we do not know how to give sufficient and
necessary qualitative conditions for identity of material objects or
persons even within one world,[46] and yet in Lewis's framework the
qualitative criteria are the only way of relating counterparts in dif-
ferent possible worlds.,

Kripke suggests that intuitively we interpret a counterfactual like
'Nixon could have lost the election' by bringing in a possible world
where Nixon lost.

> 'Possible worlds' are *stipulated*, not *discovered* by powerful
> telescopes. There is no reason why we cannot *stipulate* that,
> in talking about what would have happened to Nixon in a
> certain counterfactual situation, we are talking about what
> would have happened to *him*.[47]

It is a crucial part of Kripke's proposal that one be able to stipulate,
as a part of a description of a possible world, which individuals are
involved in it, and that this can be done not by giving a qualitative
description, but in some – yet to be specified – more direct way. Let
us try to specify it now.

---

[46] See for example Kripke, 1980, 43: "Mathematics is the only case I really
know of where [adequate necessary and sufficient conditions for identity] are
given even within a possible world, to tell the truth. I don't know of any such
conditions for identity of material objects over time, or for people. Everyone
knows what a problem this is. But, let's forget about that. What seems to be
more objectionable is that it depends on the wrong way of looking at what a
possible world is."

[47] Kripke, 1980, 44.

3.2.2 Two Versions of Haecceitism

We have just seen that in Kripke's view, it is admissible to specify a possible world by stipulating which individuals it contains. Actual-world individuals can figure directly in a description stipulating a non-actual possible world, and can, consequently, exist in more than one world. We have noted that individuals do not have to be specified only qualitatively but can also, for example, be specified using ostension or by their name. We can therefore make within Kripke's framework good sense of an individual retaining its identity while its properties change, and that holds both in the actual world and in the non-actual ones.

In our analysis of the identity puzzles concerning the ship of Theseus, and the case of Shakespeare and Bacon, we saw that a Lewisian response was to ban all changes of properties of individuals, which in effect made all of their properties essential. Kripke opposes this view. According to him, we can meaningfully ask whether certain statements concerning an actual-world individual would hold of that very same individual in a counterfactual situation where that individual's properties have changed.

Kripke admits that for most kinds of entities, we do not have available descriptions that would provide the necessary and sufficient criteria of their identity. We should then assume that there is some presumably non-descriptive fact of the matter in virtue of which we can tell whether an individual in a counterfactual situation is identical to the actual-world individual we are interested in. Nathan Salmon convincingly argues[48] that a stipulation of possible worlds in terms of actual-world individuals requires at least the adoption of haecceitism.[49] Let us have a look at why this should be the case:

Let us start by reminding ourselves of what we mean by haecceitism. Two passages from Kaplan's work are usually brought forward to define it:

> The doctrine that holds that it does make sense to ask - with-

---

[48]Salmon, 1986.

[49]It is also compatible with adoption of even more ambitious essentialist doctrines, as we shall show shortly.

out reference to common attributes or behavior - whether *this* is the same individual in another possible world, that individuals can be extended in logical space (i.e., through possible worlds)...and that a common 'thisness' may underlie extreme dissimilarity or distinct thisnesses may underlie great resemblance, I call Haecceitism.

Haecceitism holds that we can meaningfully speak of a thing itself - without reference...to individuating concepts (other than being this thing), defining qualities, essential attributes, or any other of the paraphernalia that enable us to distinguish one thing from another. It may be that each thing has essential attributes with which it is vested at all times and in each possible world in which it exists. But that is an issue posterior to whether things have trans-world being.[50]

In my view, these passages characterise two related but distinct doctrines.[51] The first passage outlines a sort of non-qualitative essentialism, according to which an individual's identity across possible worlds is warranted by a primitive thisness. The second passage suggests something more modest: that it is possible to refer to an individual without taking recourse to any particular means of identifying it. This view, at least on the face of it, is not essentialist. For the time being, I shall call the first view *full-blown haecceitism*, and the second *modest haecceitism*.

In Kripke's work, we find an analogical bifurcation.[52] On the one hand, Kripke seems to endorse a modest, non-essentialist haecceitism. The following passage suggests as much:

Philosophers...have asked, are there objects *behind* the bundle of qualities, or is the object *nothing but* the bundle? Neither is the case; this table is wooden, brown, in the room,

---

[50]Kaplan, 1975, 722-723.

[51]Salmon, in Salmon (1986), discusses various versions of haecceitism. My treatment of this topic is influenced by his views but I use the analysis for a different purpose.

[52]This issue is connected to Kripke's endorsement of essentialism but is a separate one.

> etc. It has all these properties and is not a thing without
> properties, behind them; but it should not therefore be iden-
> tified with the set, or 'bundle', of its properties, nor with
> the subset of its essential properties. Don't ask: how can I
> identify this table in another possible world, except by its
> properties? I have the table in my hands, I can point to it,
> and when I ask whether it might have been in another room,
> I am talking, by definition, about it.[53]

Adopting a somewhat Wittgensteinian tone, Kripke is trying to 'dis-
solve' an apparent problem, and to show that an object is neither a
'bundle of properties' nor anything behind it. On the other hand,
in his more essentialist mood, he says that "(roughly) *being a table*
seems to be an essential property of the table."[54]

In order to see whether we can reconcile these two views, we have
to look at the broader context of Kripke's work. By reconstructing
the role haecceitism is supposed to play, we can draw conclusions as
to which version would fit the bill.

Haecceitism? Yes, but which one?

As we saw in Kripke's criticism of Lewis, being able to describe
possible worlds by stipulating non-descriptively which individuals
are involved in them is supposed to result in a more intuitive view
of possible worlds. Adoption of transworld individuals will help us
leave behind the wrong - Lewisian – picture, according to which
possible worlds are like 'foreign countries'. Both the modest and
the full-blown version of haecceitism could do this job. We might
hold – as in full-blown haecceitism – that an haecceity is an essential
property unique to each individual, or – as in the modest version –
that haecceitism amounts to making it possible to pick out an in-
dividual without recourse to a qualitative description, but does not
amount to a commitment to an underlying essential property, haec-
ceity. In order to make his notion of transworld individual feasible,

---

[53]Kripke, 1980, 52. Compare also a passage to the same effect in Kripke, 1980,
46.

[54]Kripke, 1980, 114,footnote 57.

Kripke needs at least the modest haecceitism but his views seems compatible with full-blown haecceitism as well.

The issue of haecceitism is closely connected with the issue of identity of individuals across possible worlds – something like haecceitism is needed for the notion of transworld individuals to work. And the adoption of transworld individuals is crucial for the workings of rigid designation, because – as we know – rigid designators are supposed to denote the same referent, the same individual, in all possible worlds.[55] We can thus look at the notion of rigid designation and try to find out which version of haecceitism it necessitates.

In order to make the concept of rigid designation work, we have to be able to distinguish between the situations where there *is* a particular individual, say Bob, in some possible world, and those situations where this only *seems* to be the case. In Kripke's worlds, not everything that looks like Richard Nixon is Richard Nixon: if, for example, Richard Nixon* were, in a non-actual world, an automaton fantastically resembling the actual-world Richard Nixon, who is a human, Richard Nixon and Richard Nixon* cannot be the same individual. We can be presented with a qualitatively defined world that contains an individual closely resembling the actual-world Richard Nixon, and be asked whether such a world indeed contains Richard Nixon. And that is why even in Kripkean possible worlds we still need cross-world identity criteria. They are needed to play the role of truth-warrants, to fix the truth-values of sentences concerning individuals (e.g., Richard Nixon), thus giving the question whether Richard Nixon is in the domain of individuals of a particular possible world a determinate answer.

Kripke proposes a number of essential properties[56] that introduce some necessary conditions on an entity's identity. Some of the essential properties take the form of conditions on the constituent parts of an entity.[57] These properties provide some necessary but

---

[55]Kripke, 1980, 48.

[56]For example originating from a particular hunk of matter, and having the parents an individual in fact has.

[57]Being made of a particular hunk of matter, or being made up of atoms that have a particular atomic number are clear examples of giving identity criteria in

not sufficient criteria of identity for some kinds of entities. This, Kripke says, just reflects our state of knowledge: we do not know yet what the sufficient criteria of identity for most kinds of entities should look like.[58] Specific essential properties of the kind mentioned here cannot therefore consistently do the job of telling apart those possible worlds where an individual occurs from those where it only may seem so. And this is where haecceitism comes in. In fact, one could see it as a tool designed to do just that. Haecceitism provides the identity criteria we needed to put flesh on the concept of a transworld individual, and, indirectly, the notion of rigid designation.

Is it the modest or the full-blown haecceitism that can play this role? What we need for rigid designation to work, is necessary and sufficient criteria of cross-world identity, something an individual has in every world where it exists. Kripke proposes that in every possible world in which it exists, an entity has the property of *being the very entity it is*.[59] This is not the property of self-identity, which trivially applies to every object. On the contrary, this property is unique to each entity: for every entity x, only x has the property of *being x*.

As commonly understood,[60] a property is essential to an entity when the entity cannot fail to have it if it is to exist. It follows then that the property we have outlined in the last paragraph is an essential one. That is why we can conclude that the haecceitism Kripke needs has a essentialist import, and a weaker, modest version will not do. Kripke needs not only haecceitism, but haecceities as well.

---

terms of constituent parts. In the case of having particular parents this is less obvious, but if we take it that one's parents determine an individual's particular DNA, we can see this property as describing composition as well.

[58]Compare the passage from Kripke, 1980, 43, quoted in a footnote on p. 90.

[59]This is, I believe, how we should read the passage in Kripke, 1980, 114, footnote 57, where he says that "(roughly) *being a table* seems to be an essential property of the table."

[60]See p. 31.

3.2.3   Haecceities and Reduction

What kind of property is a Kripkean haecceity? Should we understand it as a primitive or can it be further analysed? In part, we can find a Kripkean answer in what was already stated: given our state of knowledge, and the preconditions of rigid designation, haecceities are the best we can do at the moment. Since this is co-determined by our current epistemic situation, it could change. Kripke seems to think that with respect to at least some kinds of entities, science can provide necessary identity criteria in terms of conditions on the constituent parts of those entities.[61] In other cases, Kripke admits that the very conceptual possibility of a reductive analysis is an open question:

> Although the statement that England fought Germany in 1943 perhaps cannot be reduced to any statement about individuals, nevertheless in some sense it is not a fact 'over and above' the collection of all facts about persons, and their behavior over history... Similarly, perhaps, facts about material objects are not facts 'over and above' facts about their constituent molecules... In each case we seek criteria of identity across possible worlds for certain particulars in terms of those for other, more 'basic' particulars. If statements about nations (or tribes) are not reducible to those about other more 'basic' constituents, if there is some 'open texture' in the relationship between them, we can hardly expect to give hard and fast identity criteria...[62]

Whether statements about one kind of entities are reducible to statements about another kind of entities is, at least in some cases, an open question. Its solution depends on there being bridging laws between those kinds of entities.[63] And even when an exhaustive de-

---

[61]For example in Kripke, 1980, 44, we read that "... characteristic theoretical identifications like 'Heat is the motion of molecules', are not contingent truths but necessary truths, and here of course I don't mean just physically necessary, but necessary in the highest degree – whatever that means."

[62]Kripke, 1980, 50.

[63]This is not all that needs to be said on the role of science within Kripke's framework, but we shall return to this subject shortly.

scription of an object in terms of its constituent parts is available, it may be more practical to speak in terms of nations rather than individuals when discussing history, or in terms of tables and chairs instead of molecules when ordering furniture.

In principle, Kripke might say, science may provide us one day with necessary criteria of identity for all sort of entities. But for the time being, haecceities do the job where we do not have such criteria, and are often handier even when we do have them.

Now that we saw that something like haecceities is needed to make rigid designation work, we have a somewhat better idea about the assumptions that need to be made to establish a referring relation between a designator and its referent in various possible worlds. The next question we shall try to answer is what happens in the possible worlds where the referent does not exist.

3.2.4 Persistence and Obstinacy

We have now dealt with issues connected with determining whether a referent exists in a particular possible world. In this section, we shall focus on the referring relation, in particular on the question of what happens with reference in those possible worlds where the referent does not exist.[64]

The best known definition of rigid designation tells us that a rigid designator designates the same object in every possible world,[65] In another definition we are told that a rigid designator designates the same object in every possible world in which that object exists.[66] And finally, in Kripke's letter to Kaplan, we read that "a designator D of an object x is rigid, if it designates x with respect to all possible worlds where x exists, and *never designates an object other than x with respect to any possible world.*"[67] These three definitions are

---

[64]We have given a preliminary characterisation of the notions we shall use in this section already in section 1.7.3.

[65]Kripke, 1980, 48.

[66]Kripke, 1980, 49. We introduced this definition of persistently rigid designator on p. 26.

[67]Kaplan, 1989b, 569, my italics. We used this definition earlier, on p. 47.

clearly not equivalent. Let us have a closer look at their differences.

We can safely assume that a rigid designator designates the same object at least in every world where that object exists. All three definitions make this claim. What is not clear is what happens in the worlds where the designatum does not exist.

To account for various positions with respect to this issue, Kaplan and Salmon[68] introduce some helpful distinctions. Firstly, there are rigid designators that designate the same thing in every possible world where that thing exists, and nothing in those worlds where it does not. These are called *persistent* designators. Secondly, there are rigid designators that refer to the same thing in every possible world regardless of whether their designatum exists there or not. These are the *obstinate designators*. There is yet another kind of rigid designator: *Strongly rigid designators* rigidly refer to something necessarily existent. These designators are, by definition, both persistent and obstinate. Kripke cites mathematical descriptions (such as 'the smallest prime') as examples of strongly rigid designators.[69]

The question we want to answer is: what kind of rigid designators are proper names? In the present context, we are not interested in the semantics of names of mathematical entities. We shall therefore leave them aside. Having done that, it seems quite clear that the proper names we are interested in refer to contingently existing entities. The particulars that exists in the actual world might have failed to exist, and there might have been more entities than there actually are. Had my parents never met, I would not have existed. Had they met earlier than they in fact did, I could have had an older brother. The difference between my older brother and me is that I contingently exist, while he, equally contingently, does not.

Having established that the referents we are interested in are contingently existent entities, we can conclude that proper names

---

[68]Kaplan, 1989a, and Salmon, 1982, 32-40. Salmon elaborates on a distinction proposed by Kaplan.

[69]Kripke introduces this notion in Kripke, 1980, 48. He also further distinguishes contingently existent and contingently non-existent entities from necessarily non-existing ones, like Sherlock Holmes, in Kripke, 1972a, but that is not relevant to our topic here.

are not strongly rigid designators. It remains to be seen, however, whether they are persistently or obstinately rigid.

When investigating reference in various possible worlds, it is sometimes useful to use the analogy with alternate timelines.[70] The sentence 'Anna Pilatova is dead' is false now but true any time after my death. If the name 'Anna Pilatova' did not denote anything after my demise, that is, if it were a persistent designator, the sentence above could not be true after my death. It seems thus plausible to say that the name 'Anna Pilatova' does denote someone with respect to the 22nd century, namely me. It is because the term has a denotation in that situation that the sentence 'Anna Pilatova is dead' is true with respect to this future time. We could use the same reasoning when evaluating the sentence 'It might have been the case that Anna Pilatova was never born'. Here, again, the sentence is true because there are possible worlds in which I do not exist, and in those possible worlds the name still denotes me, the actual person.

This leads to the conclusion that proper names are obstinately rigid designators. A proper name N primarily designates an actual-world entity x. Having its reference thus fixed, N shall refer to the same x also in non-actual possible worlds. In those possible worlds where x does not exist, negative existential statements involving N can still turn out to be true. This is an interesting observation because reference is often seen as a relation between a designator and its designatum, and it is in that form that reference is usually discussed in the literature. The cases where the designatum does not exist and reference to it is still successful make us re-think this very intuitive picture of reference.

Reference is especially complicated in the case of entities that are contingently non-existent in the actual world. Their reference can only be introduced by means of description (as in 'George, my older brother' or 'Jacques, the last Frenchman'), and while the proper names thus introduced are rigid, the identity of the referents is not as clear as the identity of entities that exist in the actual world.

---

[70]The following two examples, as well as the point they argue for, are adapted from Salmon, 1982, 37-39.

For example, it might turn out that the description, by which the contingently non-existent referent was introduced, fails to pick out a unique entity. It seems that in such a case there would be no fact of the matter as to who the referent is. Seemingly contingently non-existent entities could also at closer inspection turn out to be necessarily non-existent (as is claimed for fictional entities). But that depends on the treatment of fictional entities one adopts.[71] Reference to contingently existing entities differs from the reference to contingently non-existent entities because of the difference in our access to them. Our prejudice in favour of the actual may turn out to be rather a statement of sober recognition of the perils we can encounter in the realms of the merely possible.

3.2.5   Possible Worlds and Imagination

The case of the unicorn draws our attention to the limits of stipulating possible worlds by the means of specifying which entities occur in them. The fact that we can describe possible worlds this way should not make us think that anything we can imagine is genuinely possible. In Kripke's view, possible worlds are not created by our imagination: what is and is not possible is given by the modal properties of actual-world entities. We already know that in Kripke's view, necessity and contingency apply not just to statements, but to properties as well.[72]

   Modal properties are described in terms of possible worlds, so that, for example, an object has a property necessarily if it has that property in every possible world where it exists. Possible worlds were introduced to model modal properties of actual-world entities, and that is why it comes as no surprise that modal properties of

---

[71]Kripke clarifies his view of fictional entities in Kripke, 1980, 24, where he says: "So it is said that there might have been unicorns. And this is an example of something that I think is not the case. I think that even if archaeologists or geologists were to discover tomorrow some fossils conclusively showing the existence of animals in the past satisfying everything we know about unicorns from the myth of the unicorn, that would not show that there were unicorns."

[72]Kripke, 1980, 41.

actual-world entities place certain conditions on what goes on in these possible worlds. We can thus turn Kripke's analysis of necessity around and say that only those worlds are possible where the properties which are necessary in the actual world hold.

There are many kinds of necessity, which result in different kinds of commitments, and the necessity that Kripke deals with results in metaphysical commitments.[73] Kripkean necessity is about *how things are* irrespective of our knowledge, language, or context. It may well be the case that there are many more necessary properties than we shall ever know, and that these unknown properties place restrictions on possible worlds just as much as those properties we think we have already identified.[74]

Therefore, while we can imagine worlds where some properties that are necessarily in the actual world, e.g., some laws of physics, do not hold, Kripke would say that these are not really possible worlds. They are just figments of our imagination. We may call them impossible possible worlds or Oscar and Felix, but they have no place in Kripke's ontology. In Kripke's world, only those worlds are possible where everything that is necessary in the actual world holds. In this sense, Kripke is an actualist: what is possible is determined by the modal properties of the actual world. Necessary statements thus seem to play an important role in building Kripke's possible world framework. It is time we paid closer attention to them.

3.2.6   What is a priori?

As we just noted, a world's possibility is co-determined by modal properties of actual-world entities. Necessary statements play an

---

[73]The robustness of Kripke's concept of necessity is illustrated for example in Kripke, 1980, 142: "Any necessary truth, whether *a priori* or *a posteriori*, could not have turned out otherwise."

[74]On the other hand, Kripke says that "A possible world is *given by the descriptive conditions we associate with it.*" Kripke, 1980, 44, italics in the original. This seems to clash with the view which I argue Kripke adopts. The way we should see it is that in this passage Kripke emphasises the difference between Lewis's approach and his view, which he later explains in more detail.

important role here: for a world to be genuinely possible, all necessary statements must have the same truth-value in that world as they have in the actual world. Traditionally, philosophers have considered only those statements to be necessary which were known a priori. These statements were then said to be true everywhere if true at all.[75] On this traditional view, necessary statements were those that could be known by reasoning alone. The subject matter of these a priori necessary statements were certain ideas, not things in the world, which were seen as contingent.[76] It is well known that Kripke rejecta this approach, and argues that a posteriori necessary statements are not only possible, but can be known as well.

The claims concerning a posteriori necessary statements are novel, controversial, and essential to Kripke's enterprise. Before we start looking at particular statements Kripke claims are a posteriori and necessary, we shall examine what he means by a priori and a posteriori, and how he argues for the separation of the epistemic and the metaphysical distinctions.[77]

Kripke presents his clearest and least presupposition-loaded argument for the separation of the epistemic and the metaphysical distinction, as well as for the existence of a posteriori necessary statements, using an example from mathematics. Let us briefly review it.[78] Consider Goldbach's conjecture. It says that every even number greater than 2 is a sum of two primes. We do not know whether this conjecture is true. None of us therefore has any a priori knowledge in this respect. When someone finds a proof of Goldbach's conjecture and we come to believe it, it will be to us a new piece of information. It will be a posteriori evidence. Now notice, so Kripke urges, that, regardless of our ignorance, if the conjecture is true it is necessarily true, and if it is false it is necessarily false because the truth-value of a mathematical statement is not contingent. Our

---

[75]For example Duns Scotus has been interpreted as saying just that.

[76]This had a lot to do with the religious assumptions made by medieval and early modern thinkers.

[77]We have briefly introduced this subject in section 1.7.3. Here we assume the content of that section, and expand it.

[78]The argument is given in Kripke, 1980, 38-38.

ignorance of a particular statement's truth-value has no bearing on its truth-value. This is the gist of Kripke's argument.

Before we go on evaluating this argument, we have to take a closer look at Kripke's use of the terms a priori and a posteriori. This is because – as we shall see – Kripke uses those terms in a rather non-standard way.

When introducing the notion of a priori, Kripke quotes Kant's definition of it, which says that a priori truths can be known independently of experience.[79] Commenting on this definition, Kripke says that if something *can* be known a priori, we should ask ourselves for whom it is possible to know it in that way. In other words, Kripke endorses the view that a prioricity is relative to the knower. This leads him to say that

> . . . it might be best therefore, instead of using the phrase '*a priori* truth', to the extent that one uses it at all, to stick to the question of whether a particular person or knower knows something *a priori* or believes it true on basis of *a priori* evidence.[80]

We have seen above that Kripke thinks of the notion of a priori as describing the relation between a piece of information and a knower. In this quotation, however, he applies the notion to evidence. The obvious question is: What counts as a priori evidence? Kripke elucidates the concept in the following example:[81] A person who works with a computer knows that the computer can answer whether a particular number is prime. No person has calculated that this particular number is prime, but the machine gave us the answer. If we then believe that this particular number is indeed prime, we believe it on the basis of our knowledge of the capacities of the computer. It seems therefore that we believe it on a posteriori grounds. Nonetheless, someone who made the requisite calculations himself could believe a priori that the number in question is prime. But

---

[79]Kripke, 1980, 34.
[80]Kripke, 1980, 35.
[81]Kripke, 1980, 35.

that does not mean that person is necessarily right - one can make mistakes in a priori reasoning.

> Something can be known, or at least rationally believed, *a priori*, without being quite certain. You've read a proof in the math book; and, though you think it's correct, maybe you've made a mistake... You've made a computation, perhaps with an error.[82]

It seems that according to Kripke, a belief is a priori if it is based on a priori evidence, and evidence is a priori if it is available to the agent without recourse to external fact checking. This notion of a prioricity is inherently context-relative. Something that is a priori for one person need not be so for another one. This notion is very different from that which was used by Kant.

The underlying reason for the divergence between the traditional notion of a priori and Kripke's notion may well lie in Kripke's approach to analyticity, and indirectly, his approach to necessity. In the Kantian tradition, a statement is analytic if its truth can be determined by analysis of the terms involved alone. The necessity associated with a prioricity is thus intended to be primarily of an epistemic, but secondarily also of a semantic kind, derived from the meaning of terms involved in a sentence. The semantic kind of necessity, i.e., analyticity, was famously attacked by Quine[83] and has largely fallen into disrepute. Kripke does not use that notion. But in the traditional Kantian picture, there was a connection between analyticity and the a priori: a statement was analytically true if its truth could be known just by analysis of the meaning of the concepts involved, and the knowledge thus derived was then a priori. This sort of 'semantic necessity' is not available to Kripke, who ascribes necessity to facts about the world, and the statements that express them.

In the Kantian picture, a competent speaker can figure out that some statements are necessary in of virtue of his command of language. In the Kripkean picture, there is no parallel to this: different

---

[82]Kripke, 1980, 39.
[83]In Quine, 1961b.

agents know different things about the world, and that is why a priori knowledge is speaker-relative. The shift of meaning of 'a priori' between Kant and Kripke is to a large extent a result of the shift of interest from semantics to metaphysics. In this light, it is hardly surprising that Kripke's notion of a priori is rather thin. Because is not connected to necessity, it does not guarantee knowledge, and because it is not connected to a competence that is generally shared between agents (i.e., language), it is speaker-relative.

3.2.7   Natural Kinds and Haecceities

When dealing with Kripke's approach to necessity and the a priori, one cannot avoid an overview of the kinds of statements, which, as Kripke claims, can a priori be known to be necessarily true if true at all. And any such overview has to include Kripke's analysis of natural kind terms. Though natural kind terms are not a subject of our primary interest, the connection between haecceities and reference is, which more than justifies our little excursion into natural kinds.[84]

We shall not speculate about what he would have or should have said had he elaborated more on the topic of natural kinds. In particular, we shall not assume that any gaps one may find in Kripke's views can be filled by Putnam's views.[85]

---

[84]We shall deal with natural kinds in a rather cursory manner. However, many of the authors whose work on proper names we mentioned or used in our explorations have also written on natural kinds. It was mainly the work of Kripke (1980), and Putnam (1975a) that started the debate on natural kinds, but important contribution were made also by Wiggins (1980), Burge (1973), McGinn (1976), and Salmon (1982), as well as all those whose articles were collected in French, Uehling, and Wettstein (1986).

[85]Kripke says that the views on natural kinds and substances he entertains in *Naming and Necessity* (1980) have many points of contact with Putnam's pre-1963 writings, but also that "there are some divergences between Putnam's approach and mine."(Kripke, 1980, 122, footnote 62).

Kinds Are Introduced

Kripke deals with natural kinds mainly in his *Naming and Necessity*,[86] focusing above all on the issues of necessary and contingent properties. His aim is to refute the view that observable properties usually associated with natural kind individuals and samples – e.g., *being tawny yellow with blackish transverse stripes and white belly* in the case of tigers – are necessary of them. We could, Kripke says, imagine circumstances where none of these commonly associated properties apply to a particular individual, and yet these would still be circumstances where that individual still belongs to the kind that is thus characterised. The reverse also holds: something might have all the identifying characteristics of some natural kind, yet form a separate kind.[87] Therefore, Kripke concludes, "possession of most of these properties [by which we originally identified a kind] need not be a necessary condition for membership in the kind, nor need it be a sufficient condition."[88] Contrary to Frege's, Russell's, and even Mill's views, natural kind terms are not descriptive. According to Kripke, they are much more like singular terms than has been thought, and his treatment of them is largely parallel to his treatment of proper names.

In Kripke's view, proper names are attached to individuals by a hypothetical 'baptism', where a description, ostension or both can be used. Natural kinds also undergo a sort of 'baptism' where a definition and/or ostension is used,[89] as in "Gold is the substance

---

[86]Especially in Lecture III of Kripke, 1980.

[87]See Kripke, 1980, 119, and the following: "Even though we don't know the internal structure of tigers, we suppose - and let us suppose that we are right - that tigers form a certain species or natural kind. We can then imagine that there should be a creature which, though having all the external appearance of tigers, differs from them internally enough so that we should say that it is not the same kind of thing. We can imagine it without knowing anything about this internal structure. We can say in advance that we use the term 'tiger' to designate a species, and that anything not of this species, even though it looks like a tiger, is not in fact a tiger." (Kripke, 1980, 121)

[88]Kripke, 1980, 121.

[89]We can, just like in the case of proper names, imagine circumstances where a natural kind is baptised using a definition only. For example, a number of

instantiated by the items over here, or at any rate, by almost all of them."[90] This sort of definition, which is used to introduce a natural kind term, does not express a necessary truth because the natural kind would have existed even if any particular items did not. Kripke says that

> ...in general, terms for natural kinds (e.g., animal, vegetable, and chemical kinds) get their reference fixed in this way [by the above mentioned sort of definition]; the substance is defined as the kind instantiated by (almost all of) a given sample.[91]

The reference of natural kind terms is rather similar to the reference of proper names. Natural kind terms are not descriptive: their referents may fail to have any of the properties commonly attributed to them, and something may have all of those properties and not be the intended referent. The descriptions used in fixing the referent also do not form a part of their meaning. In Kripke's view, natural kind terms are rigid designators that designate directly.

On Being A Natural Kind

Natural kind terms are unlike proper names in that the question of the ontological status of their referents is somewhat more complicated. Everyone has some idea of what an individual is but few people outside of philosophy have ever heard about natural kinds. And it is difficult to see what ontological status Kripke intends natural kinds to have but he does give us some clues by saying that "the original concept of cat is: *that kind of thing*, where the kind can be identified by paradigmatic instances."[92] He also suggests that a natural kind can be identified *by* its instances but not *with* them – a change in their number does not amount to any change in the kind itself. Natural kinds must therefore be to some degree ontologically

---

unstable chemical elements high in the periodic table were defined by their atomic number years before they were first synthesised.

[90]Kripke, 1980, 135.

[91]Kripke, 1980, 135-136.

[92]Kripke, 1980, 122.

independent of their instances, which means that they are abstract entities.

What all instances of a natural kind have to share is their essential properties. In Kripke's view,[93] it is the task of science to discover what those essential properties are. We can, however, try to reconstruct a minimal picture of the semantics of natural kind terms without making any claims about the essential properties of particular kinds.

In general, we can assume – in parallel with individuals – that natural kinds have haecceities, which could be expressed as *being this sort of thing*.[94] We can further suppose that instances of natural kinds have also individual haecceities (of the form *being this very thing* or *being this very individual*). Indeed, unless we assume that we have identified the essential properties for every natural kind, it seems that we need to assume natural kind haecceities in order to make rigid designation work. The reasoning behind this is the same we used for proper names. In the case of natural kinds that are exemplified by individuals (like 'tiger') or samples (like 'gold'), if we want to be able to refer rigidly to an individual or a sample, we have to assume also individual haecceities.[95]

Kripke does not say anything explicitly about the relation between individual haecceities and the haecceities of kinds. However, we can recall his treatment of the Nixon case,[96] where it is said that if Nixon is human, then he is necessarily so. This indicates that there is a (metaphysical) necessity relating the two kinds of haecceities. Addressing the question of mutual relations between various natural kind haecceities, Kripke says that "of many such statements [like 'Cats are animals'], especially those subsuming one species under another, we know a priori that, if they are true at all, they are necessarily true."[97] This claim relies on further assumptions related

---

[93]At for example Kripke, 1980, 138

[94]See the above quoted passage Kripke, 1980, 122.

[95]In the case of samples this point is somewhat stretched but we can imagine a situation where we want to refer to a particular piece of gold.

[96]Kripke, 1980, 46.

[97]Kripke, 1980, 138.

not only to the ontology of natural kinds but also to the problematic notion of a priori. We shall analyse the presuppositions of this sort of claims in the following section.

Hierarchies of Haecceities, Tony The Tiger, and Other Beasts

We can get a better idea of Kripke's theory of natural kinds if we apply the aforesaid to a couple of examples. This should help us organise what we have learned so far. Let us therefore conduct an analysis of the epistemological and the modal status of the following sentences using a Kripkean perspective.

(1) Dogs bark.

(2) Tony is a tiger.

(3) Dogs are mammals.

If we read sentence (1) as a generic statement, it turns out to be contingent and a posteriori in virtue of the non-descriptionality of natural kind terms. A natural kind term refers directly and rigidly to the kind, and the kind refers directly to its instances. Strictly speaking, this needs to be qualified: we cannot say that a natural kind term rigidly refers to its instances because its extension can vary.[98] In this generic statement, 'barking' is predicated of items belonging to the natural kind 'dog'. On a naive reading, a single instance of a non-barking dog should falsify the sentence, and that is clearly undesirable.[99] Kripke does not deal in his work with the problems raised by this kind of generic sentences.

We can analyse sentence (2) as necessary a posteriori by analogy with the Nixon example. As we pointed out, belonging to a particular kind is supposed to be a part of an individual's haecceity.

---

[98]In a non-actual possible world, the extension of a natural kind term may be different than in the actual world. If the definition of rigid designation demands that a rigid designator refers to the same thing in every possible world where that thing exists, we would have to amend the definition to account for the designation of natural kind terms.

[99]For more on the problems of non-barking dogs, see Carlson, 1977, 56nn.

We can read (2) as claiming that *being a tiger* is a feature of Tony's haecceity. Tony's natural kind haecceity, *being this sort of thing*, puts him in the extension of the kind 'tiger', which presumably has its own haecceity (it is hard to tell what that would be except that clearly the kind 'tiger' is not a tiger).[100]

Intuitively, a Kripkean analysis of the predicative use of natural kind terms – as seen here – is not very satisfactory. The non-descriptionality of natural kind terms together with the metaphysical necessity attendant to the relation between an individual and the kind to which it belongs, makes it somewhat unclear why sentences like (2) are used to convey information. There does not seem to be a connection between the semantics Kripke predicts for this sort of sentence and the function it has in a real discourse.

Sentence (3) should be analysed as necessary because subsumption of one species under another, if true, is necessary.[101] Moreover, we should know a priori that if such subsumption is true, it is necessarily true. Intuitively, this is not a very satisfactory position, because it ascribes metaphysical necessity to taxonomic principles. Yet we know that taxonomy changes all the time. A Kripkean analysis of sentence (3) makes ambitious realist claims about taxonomy. These are not trivial assumptions to make, and one would expect Kripke to provide some sort of argument for this position but, unfortunately, this he does not do.

Any reconstruction of Kripke's analysis of the semantics of natural kind terms is tentative because there is little textual basis to work with. It is not hard to see why the analogy between natural kind terms in their referential function and singular terms is tempting. However, once we analyse natural kind terms as non-descriptional rigid designators, we run into problems with their predicative use. In his treatment of natural kinds, Kripke does not deal with the predicative use of natural kind terms. Using an extrapolation from the Nixon case is safe but the resulting position is not satisfactory.

---

[100]See Frege, 1893, re-print in English 1952.
[101]See a quotation to this effect on p. 108.

It is surprising that Kripke's views on natural kinds, which have been very influential, are so incomplete with respect to even rather obvious issues, such as those we mentioned above.

3.2.8 Necessary Statements and Their Commitments

When investigating the interface between semantics, metaphysics, and the role of science in Kripke's work, we quickly find that the borders between the three shift depending on what kind of statements and entities we consider. In this section, we shall look at different kinds of a posteriori necessary statements, and try to find out what has to be presupposed if we are to analyse them in a Kripkean fashion.

Let us start by considering statements that express identity between individuals, like 'Hesperus is Phosphorus', 'Cicero is Tully', and 'I am Anna Pilatova'. According to Kripke, these statements – that is, statements composed of two distinct singular terms, each of which is either a proper name or an indexical – are a posteriori and necessary.[102]

That these statements should be a posteriori seems immediately plausible: that Hesperus and Phosphorus refer to the same planet was at one time a discovery; that Cicero is Tully, or Woody Allen is Alexander Konigsberg is still news to some. And while I may know that I am Anna Pilatova a priori, to others it is a posteriori, otherwise there would be no need for me to ever introduce myself. And if ever I suffered amnesia, I would be happy to find out what my name is.[103]

The a posteriori status of these identity statements being settled, let us turn to the claim that they are necessary. Where does that necessity come from?

---

[102]See his argument at Kripke, 1980, 101-105.

[103]This issue borders on the problems connected with describing what it takes to know who one is. If all I forget is my name (which is then supplied to me), the situation is different from one of forgetting everything except my name (then knowing that I am Anna Pilatova is not very helpful).

Once we treat proper names (and indexicals) as obstinately rigid designators, we assume that they designate the same thing in all possible worlds regardless of the existence of the referent in any given world, with the possible exception of the actual one.[104] .Coreferential singular terms designate the same thing in every possible world, which is by definition equivalent to saying that if two or more singular terms are coreferential, they are necessarily so. And, consequently, the identity statement that connects the two will be necessarily true as well. But even if singular terms were only persistent designators, and for example 'Hesperus' and 'Phosphorus' were coreferential names referring to Venus, an identity statement like 'Hesperus is Phosphorus' would still be true in every world where Venus exists. The sentence 'If Hesperus exists, it is identical with Phosphorus' would be never be false. Therefore, even if singular terms were only persistent designators, we would still get necessary identity statements that make an assumption of existence explicit, and that can be taken as close enough to necessity *tout court*.[105]

In order to arrive at this conclusion, we had to assume that singular terms are rigid designators, and in order to make that claim, we had to presuppose at least an essentialist version of haecceitism. That is, we had to assume that in every world where it exists, the referent of a rigidly designating expression has the property of *being the very entity it is*.

Kripke, however, seems to argue that the necessity of certain kinds of identity statements must be derived by 'philosophical analysis':

> Certain statements - and the identity statement is a paradigm
> of such a statement on my view - if true at all, must be neces-
> sarily true. One does know *a priori*, by philosophical analy-
> sis, that *if* such an identity statement is true it is necessarily

---

[104]If an obstinately rigid designator has no referent in the actual world, the situation becomes more complex. It could turn out that the entity is for example fictional.

[105]There are well-known epistemic problems that arise in connection with identity statements. Kripke pointed these out in his puzzle about Pierre (Kripke, 1979). We deal with this topic at length in section 4.2.

true.[106]

It is unclear whether the 'philosophical analysis' here mentioned is supposed to be a semantic analysis, in which case this argument would be compatible with the reasoning we outlined above, or whether the analysis Kripke intends here is supposed to amount to something more.

I suspect the above-mentioned 'philosophical analysis' must mean more than an analysis of language, because Kripke seems to use this reasoning to support his analysis of identity statements involving natural kinds and substances. Kripke claims that statements such as 'Water is H_2O', 'Gold is the element with the atomic number 79,[107] or 'Cats are mammals' are necessary, and, moreover, if they are true then they are a priori so, and can a priori be known to be necessary.

The kind of essentialism inherent in Kripke's claim that these statements expressing 'theoretical identifications'[108] are necessary goes beyond the adoption of the ontology of natural kinds and their haecceities. The modal status of these statements suggests concrete essentialist principles concerning natural kinds, such as that subatomic composition is a necessary feature of an atomic element, and that being a subkind of a taxonomically higher kind is an essential feature of a species. This is essentialism with much more metaphysical bite than the one we needed to make the rigid designation of proper names and natural kind terms work.

In addition, it seems now plausible to interpret the quotation above as an argument saying that for some sorts of theoretical identifications, we know a priori that if they are true, then they are necessarily so. We may be wrong about the particulars, it may just conceivably turn out that an atom of gold has 78 protons, but the principle that an identification of an element with a particular atomic number if true must necessarily be true is supposed to be known a priori.

---

[106]Kripke, 1980, 109.

[107]See Kripke, 1980, 116-117.

[108]As he calls them e.g., in Kripke, 1980, 116.

In parallel to the above-mentioned necessary statements about natural kinds, Kripke lists various a posteriori necessary statements about individuals. Statements like 'Richard Nixon is not an inanimate object', 'Elisabeth II sprang from the very gametes she actually sprang from', and 'This table is necessarily made of the very hunk of wood it is actually made of'[109] suggest various essentialist principles concerning individuals.

Nathan Salmon[110] convincingly argues that the essentialist principles implied by these kinds of statements (e.g., 'Gold is an element with the atomic number 79', and 'Nixon is not an inanimate object') cannot be derived from the semantics of the terms alone, or possibly from their semantics and some uncontroversial premises. The necessity of these statements derives from a previously assumed metaphysical theory of essentialism that is independent of a theory of reference.

It seems thus that Kripke makes two kinds of claims regarding essences. Firstly, he claims that certain general essentialist principles are known a priori.[111] An example of this kind of principles is that *having the parents one actually has* is an essential property of persons. Secondly, Kripke claims that some particular a posteriori statements, e.g., 'Gold is the element with atomic number 79', if true at all, are necessarily true. Neither of these positions can be derived from an analysis of language. Let us now have a closer look at the assumptions that allow Kripke to make these claims.

3.2.9   Kripke's Scientific Realism

The kind of essentialism involved in statements such as 'Gold is the element with the atomic number 79' involves claims about *how things are in the world*, which makes it a claim about metaphysical necessity, as Kripke freely admits. In this section, we shall look at Kripke's motivation for accepting this view, and say something

---

[109]In all these cases, the property mentioned is supposed to be essential (necessary) of that particular individual or thing.

[110]Salmon, 1982.

[111]We shall investigate this claim in more detail in the following section.

about the assumptions that need to be made in order to derive it.

As we already indicated, Kripke claims that it is known a priori of some characteristic theoretical identifications (such as 'Water is H_2O') that if they are true, then they are necessarily so. This presupposes that we know certain essentialist principles in their general form (e.g., 'chemical composition is a necessary property of a chemical compound') in some sort of a priori manner. However, it is very hard to see what kind of a prioricity Kripke has in mind here.[112] It is quite implausible to assume that we could derive these essentialist principles by reflection on the meaning of natural kind terms alone. We must therefore assume that there is yet another kind of considerations at play here. Just what they are becomes clearer once we look at claims of an even more radical kind, which Kripke also seems to endorse, namely that some essentialist claims – and not just principles – concerning particular natural kinds are metaphysically necessary.

The motivation behind these claims emerges in passages such as:

> Such statements [as 'Gold is an element with the atomic number 79'] representing scientific discoveries about what this stuff *is*, are not contingent truths but necessary truths in the strictest possible sense.[113]

What this passage[114] illustrates, is Kripke's belief that science can discover not only truths about the actual world – which would be contingent – but also truths about all possible worlds, that is, necessary or essential truths.

Characteristic theoretical identifications such as 'Heat is the motion of molecules' might be interpreted as definitions, hence as analytic, a priori, and without any essentialist import. Yet Kripke insists that this is not the reading he intends, stating that

---

[112]The lack of clarity in Kripke's notion of a priori, which we investigated in section 3.2.6, makes things only worse.

[113]Kripke, 1980, 125.

[114]And passages to the same effect, such as "...whether science can discover empirically that certain properties are properties are necessary of cows, or of tigers, is another question, which I answer affirmatively." (Kripke, 1980, 128)

> . . . characteristic theoretical identifications like 'Heat is the
> motion of molecules', are not contingent truths but necessary
> truths, and here of course I don't mean just physically nec-
> essary, but necessary in the highest degree - whatever that
> means.[115]

In Kripke's view, science is quite clearly assumed to be the tool of
discovering how things are in a metaphysical sense, that is, inde-
pendent of our perception of them, our language or knowledge. The
meaning of the terms that various sciences use, like 'heat', 'light'
or 'tiger', is taken to be independent of changes in our knowledge.
Their meaning is supposed to be constant because they represent
kinds, unchangeable abstract entities, and while the extension of a
natural kind may change, its essence does not – and science is seen
as the right tool for uncovering that essence.

> Note that on the present view, scientific discoveries of species
> essence do not constitute a 'change of meaning'; the possibil-
> ity of such discoveries was part of the original enterprise.[116]

These are some very strong assumptions, which should be backed
by some powerful arguments. Kripke, however, does not seem to
argue for his position at all. These basic presuppositions are just
taken for granted. Most, if not all, of the a posteriori necessary
claims Kripke makes are supported by Kripke's view's on science,
and cannot not be upheld without making metaphysical assump-
tions. Kripke's views on science are not, and cannot be motivated
by the semantics of natural language – metaphysics is inherently in-
dependent of language. Our interest is in the semantic of natural
language, and that is why we shall try to limit to a minimum the
amount of assumptions that cannot be justified by an analysis of
language. In this enterprise, metaphysics is something we should
try to do without.[117]

---

[115]Kripke, 1980, 99.

[116]Kripke, 1980, 138.

[117]I think one should be cautious with metaphysics in general. It may some-
times be useful to introduce into one's theory posits that are not further ex-
plained but they should be treated as such.

3.2.10   Kripke Without Metaphysical Assumptions

The question remains what Kripke's position would look like without scientific realism, that is, without the view that science discovers metaphysically necessary truths. If he were not a scientific realist, could he still claim that proper names and natural kind terms are rigid designators?

It turns out that he could very well do without scientific realism. The essentialism necessitated by rigid designation of proper names or natural kind terms only requires the essential properties of the form *being the very entity it is* for individuals, and *being this sort of thing* for natural kind terms, that is, haecceities.[118] In the case of natural kind terms, we adopted into our ontology natural kinds as abstract entities. If we are willing to make these assumptions, we can have rigid designation for both proper names and natural kind terms, a position that is essentialist in terms of further unanalysed haecceities. Furthermore, haecceities can be treated as posits simply introduced to make rigid designation work. And rigid designation is a claim about the semantics of singular and natural kinds terms. While it is open to refutation by counterexamples from spoken language, we have not yet encountered an example arguing directly against it.

We should note that the definition of rigid designation does not imply that once it is adopted, it has to be assumed for all the kinds of terms Kripke ascribes it to. We could hold that while proper names are rigid designators, natural kind terms are not, and analyse characteristic theoretical identifications not as referring to natural kinds but rather treat them as definitions. Depending on evidence from language, we can decide for what kinds of entities we want to assume haecceities. The rigid designation thesis implies some ontological commitments - adoption of haecceities - and we can decide for each kind of individuals whether to take that step or not.

---

[118]The parallel between the haecceitism assumed in rigid designation of proper names and that of natural kind terms is not explored by Salmon, 1982 but it is, I believe, compatible with his views.

The story is different for the more particular essentialist principles because they go beyond the assumption of haecceitism. Kripke's motivation of these principles is twofold: in their general form (e.g., '*being made of a particular chunk of material* as an essential property of an artefact'), these principles are supposed to be accessible to us a priori, by philosophical reflection.[119] The problem with this sort motivation is that we can easily imagine a situation where someone arrives by her philosophical reflection at different conclusions. It is hard to imagine how one could settle differences of opinion in the realm of the a priori. Motivation from a priori reflection thus should be treated as idiosyncratic. Among other things, it is not supported by generally shared intuition about the meaning of the relevant words in language, and, as we pointed out in section 1.3, while we have to rely on our semantic intuitions to some degree, where this clearly gives rise to disagreement, the claim has to be supported by an independent argument.

So, Kripke's claim that particular essentialist statements (e.g., 'Gold is an element with the atomic number 79') are metaphysically necessary should be seen as motivated by strictly and exclusively by scientific realism. As we already pointed out, Kripke does not argue for scientific realism, he simply assumes a particular form of it. We find it objectionable to employ controversial metaphysical assumptions in the context of a study of the semantics of natural language. We may be unable to do semantics of modal statements without employing our pre-philosophical modal intuitions, in case those modal intuitions vary among speakers and, on top of that, are not supported by generally shared semantic intuitions, one should seek some other kind of argument. Kripke's reliance on scientific realism (without a convincing defence of that position) in the course of a semantic analysis is an example of an idiosyncratic use of modal intuitions, and as such it should be avoided.

When mentioning modal intuitions earlier (on p. 15), we said that we should not presuppose any connection between the semantic and the modal intuitions. Now we see that some connection should

---

[119]See the above-quoted passage of Kripke, 1980, 109.

be assumed. Where modal intuitions are not supported by semantic ones, or contradict them, they become problematic.

But what would happen with Kripke's analysis of modality without scientific realism? As we have shown, what is possible or necessary is given by the modal properties of actual-world entities. If we do not know how to find out what the necessary properties of actual-world entities are, we are equally at loss when trying to figure out whether any possible world is indeed possible. Remember that whether something is or is not possible is independent of our ability to imagine such a state of affairs. Possible worlds that have to comply with metaphysical necessities – and Kripkean possible worlds are of this kind – are a liability if we cannot find out what those necessities are.[120] If we are to give up metaphysics, we have to have a story about the kind of possible worlds we do want to work with, and about the necessity that holds such a framework together.

3.2.11 Conclusion

Dealing with Kripke's conception of the semantics of modal statements, we started off by reconstructing his version of possible worlds, focusing on the claim that it is admissible to stipulate directly which individuals are involved in a possible world. According to Kripke, one may stipulate that an actual-world individual is involved in a non-actual possible world. This is equivalent to the adoption of transworld individuals. These, in turn, are needed if we want to make sense of the notion of rigid designation - we saw in our analysis of Lewis's work that in a framework where individuals are worldbound the concept of rigid designation made no sense.

The main problem with transworld individuals is that one has to have a way of telling what counts as the same individual across various possible worlds, that is, under what conditions an individual retains its identity under change. In other words, one needs to be

---

[120]Our approach to the question of necessary properties may, however, differ depending on the kind of entity we consider. We may well take a piecemeal approach to ontology. We could believe that mathematical entities have necessary properties while rejecting non-trivial essentialism concerning persons.

able to tell when something that may look quite unlike a particular actual-world individual is still our actual-world individual, and *vice versa*, when something that looks just like an actual-world individual is something else. This job cannot be done by anything less than sufficient and necessary identity criteria for individuals, and haecceities are the minimal presupposition one has to make. Particular essential properties (like essentiality of origin) can not do the job because they provide the necessary, but not the sufficient identity criteria. This is why the haecceities were brought in. A haecceity, as we saw, is a property that an individual has to have in every possible world in which it exists. It is the property of *being the very entity it is*. We can either treat it as a primitive (which, as Kripke says, is usually the practical thing to do) or analyse it further. The possibility of further analysing haecceities depends on there being bridging laws that would facilitate a reduction of identity criteria for one kind of entities to the identity criteria of other constitutive entities. The viability of ontological reduction of essential properties for particular kinds of entities is, Kripke says, to be decided by scientific inquiry.

The semantics Kripke proposes for natural kind terms is constructed to a degree in parallel with the semantics of proper names. The range of problems presented by the reference of natural kind terms, however, is significantly different from the problems of reference of singular terms,[121] and in the present work we touched upon them only in a cursory way.

The main problem with Kripke's possible world framework as we found it had to do with the role of metaphysical necessity, which was brought in to determine which worlds are possible and which only seem to be so. Metaphysical necessity is inherently epistemically opaque, which means that we – both as speakers or a linguistic community in general – are not in the position to know which worlds are possible and which are not, and this is built into the very core of Kripke's possible world framework.

---

[121]For an exposition of some of the problems see Carlson, 1977, Salmon, 1982, or Lowe, 1997.

Moreover, Kripke investigates metaphysical necessity while neglecting other kinds of necessity. In natural language, modal terms are used in a variety of ways, and a possible world framework should have enough flexibility to capture them. In natural language, we seem to use successfully various kinds of modal statements, and we are, at least sometimes, fairly certain of their truth or falsity (which is hardly ever the case with metaphysical necessity). What we want is a framework that would enable us to better model how modal statements function in a natural discourse. And describing a possible world framework that takes its inspiration from real discourse shall be the task of the rest of this chapter.

3.3 Stalnaker's Worlds

In our analysis of Lewis's approach to possible worlds, we concluded that the realism that characterises it necessitates the adoption of worldbound individuals, which in turn leads to the counterpart theory. Both of these features detract from the theory's potential usefulness in an analysis of natural language. Kripke's approach does not suffer from these particular drawbacks but the elements of metaphysics, which could not be justified from the semantics of natural language, make his framework less than an ideal candidate for a theory we would like to work with. Stalnaker's framework seems a more attractive option – while it works with transworld individuals, metaphysical assumptions seem to be absent from it. That is why in the following sections, we shall investigate Stalnaker's approach to possible worlds in detail: we shall look at his motivation for using possible worlds, the requirements he sets for a possible world framework, and the way he goes about building one. An analysis of the problems he encounters, and of the way he deals with them, should help us to gather useful hints for our own analysis of modality.

Stalnaker has been influenced by the work of David Lewis.[122] We analysed some parts of Lewis's work at the beginning of this chap-

---

[122]See for example Stalnaker's contributions in Harper, Stalnaker, and Pearce, 1981.

ter, and concluded that its use for an analysis of natural language is problematic. That is why here we shall focus on those parts of Stalnaker's work where he departs, directly or indirectly, from Lewis's line of thinking.

3.3.1   Stalnaker's Motivation

In order to understand a possible world framework, one should first look at the problems it is supposed to help solve. In the present case, this is not too difficult. Although Stalnaker's work covers a large number of topics, its main orientation remained constant: his primary interest has been in understanding representational mental states and their role in explaining human behaviour.

> Representational mental states should be understood primarily in terms of the role they play in the characterization and explanation of action. What is essential to rational action is that the agent be confronted with a range of alternative possible outcomes of some alternative possible actions. The agent has attitudes, pro and con, toward the different possible outcomes, and beliefs about the contribution which the alternative actions would make to determining the outcome.[123]

In this picture, people are seen primarily as agents, and language is treated as just one of the means by which an agent's beliefs can be manipulated in a systematic fashion. Stalnaker's main goal is to explain the connections between representational mental states and actions. He is interested in the attitudes, e.g., beliefs, which agents entertain with respect to alternative outcomes of their actions, i.e., with respect to possible states of the world. Which possible states of the world are relevant to an action, both as starting conditions and as results, is determined by the context. Propositions are then introduced as a way of distinguishing between relevant alternatives, and agents' attitudes are treated in terms of their attitudes to propositions.[124]

---

[123]Stalnaker, 1984, 4.

[124]This characterisation is in line with Stalnaker's own description of his aims as he presents it for example in Guttenplan, 1995, 561-568.

Stalnaker aims at balancing, in his account, two basic aspects of
an explanation of behaviour. On the one hand, one can explain be-
haviour in terms of agents' attitudes and their perceptions of what
the different possible outcomes of their actions are. On the other
hand, one must take into account that if agents' actions are to pro-
duce, in general, the desired effect, they must be guided by a more
or less accurate model of reality. To put it differently, we have on
the one hand the perspective of the agent, and on the other hand
the world she encounters. In order to describe the beliefs and desires
that guide and motivate the agent in her actions, we have to take
into account her epistemic situation. On the other hand, to provide
a full description of an action, and of the success the agent aims at,
we have to bring into the picture the way the world is.[125] For exam-
ple, if I want to buy some lettuce for a salad, I succeed if I buy not
just something I believe to be lettuce, but only if that thing is, in
fact, lettuce. In a description of intentional behaviour, it is crucial
to make the right sort of connection between the epistemic possible
worlds of the agent and the possible worlds that are independent of
an agent's wishes, beliefs or desires. Developing a framework that
successfully balances out these two perspectives is, I think, what
Stalnaker takes to be the core task of most of his work.

Our main aim here is to put flesh on the idea of this balance.
It may seem that the two desiderata of explanation of behaviour
sketched above – i.e., basically the internalist and the externalist
perspective – lead to two distinct notions of content. On the one
hand, in order to explain why an agent acts the way she does, we
can ask about her perception of the situation she finds herself in. In
characterising it, we feel compelled to employ the narrow, internalist
conception of content. On the other hand, in order explain why
the agent can act with a degree of certainty based on her previous
experience regarding the outcome of her actions, we need a notion of
content that takes into account the contribution of the environment,
that is, the broad, externalist content. These two notions of content
rest on different notions of possible worlds. We shall try to see how

---

[125]See Stalnaker, 1984, 18-20.

Stalnaker balances out the internalist and the externalist perspective and the notions of possible worlds implied by them. To do this, we shall first look more closely at the main concepts Stalnaker uses, and explore their connections.

3.3.2   Possible Worlds

Unlike Kripke, and like Lewis, Stalnaker is explicit about the kind of possible worlds he wants to work with. It is possible, however, that his notion of possible worlds, even if seemingly explicit, hides tensions resulting from trying to account for both narrow and broad content. The conception of possible worlds that a theory uses, has extensive consequences for the notions of proposition and content, and that is why we shall focus on detecting possible tension between narrow and broad content throughout out inquiry.

Stalnaker deals explicitly with the concept of possible worlds in his oft-quoted article 'Possible Worlds',[126] which later became a part of his *Inquiry*.[127] He outlines his own conception of possible world by contrasting the differences between it and Lewis's notion of possible worlds. We shall look now at how Stalnaker understands Lewis's framework, at the points of difference between the two authors, and at the final picture arising from the comparison.

Stalnaker characterises Lewis's theory of possible worlds as adherence to the four following theses:[128]

> (1) Possible worlds exist.

> (2) Other possible worlds are things of the same kind as the actual world - 'I and my surroundings.'[129] Our actual world

---

[126]Stalnaker, 1979, 225-235.

[127]Stalnaker, 1984, Chapter3.

[128]I have already used Stalnaker's characterisation of Lewis's theory as Stalnaker presents it in Stalnaker, 1979, 227 in the sections of this chapter pertaining to Lewis. It is, however, in the interest of easier reading that they should be repeated here.

[129]This quotation, as well as other quotations and direct paraphrases of Lewis's work in this section come from and are based on Lewis, 1979a. The quotations in italics come from Stalnaker, 1979.

is only one world among others. We call it actual not because
it differs in kind from all the rest but because it is the world
we inhabit.

(3) The indexical analysis of 'actual' is the correct analysis.
"'Actual' is an indexical. It depends for its reference on
the circumstances of utterance, to wit the world where the
utterance is located."

(4) Possible worlds cannot be reduced to anything more ba-
sic. Lewis says, "When I profess realism about possible
worlds, I mean to be taken literally. Possible worlds are
what they are and not some other thing."

The first thesis, Stalnaker says, is compatible with Lewis's claim that
we believe in possible worlds in virtue of believing that things might
have been different from the way they are.[130] Lewis claims that
'ways things might have been' exist but says so far nothing about
the nature of those 'ways'. It is the second thesis that contains the
ontological commitment. Here Lewis says that possible worlds are
concrete particulars. The actual world is 'I and my surroundings',
and other possible worlds are more things like that. However, the
motivation Lewis gives for thesis (1) – the claim that possible worlds
are 'ways things might have been' – does not support thesis (2),
which states that worlds are concrete particulars. If possible worlds
are 'ways things might have been', then the actual world should be
'the way things are' rather than 'I and my surroundings'. 'The way
things are' is a property of a world, not a world itself. This is an im-
portant distinction because a property can exist unsubstantiated[131],
and *the way the world is* could therefore exist even if the world that
would be that way did not. Can we conclude that (2) is based on
an equivocation between 'the actual world' in the sense given by 'I
and my surroundings', and the sense given by 'the way things are'?
Yes, but (2) also has a deeper motivation that relates to thesis (3).

---

[130]The reasoning in this paragraph closely follows Stalnaker's reasoning in Stal-
naker, 1979, 227-228.

[131]This is true of most properties, and all the properties we are interested in
at the moment. It does not hold for properties like *being this person*.

Thesis (3) tells us that *being actual* is a world-relative attribute. According to Lewis, the actual world is special because it alone is the concrete world. But this is a contingent fact: from the viewpoint of a counterfactual possible world, that world would be actual.[132] At this point, Stalnaker starts disagreeing with Lewis. He asks

> ...if there is no absolute property of actuality, does this not mean that, looking at things from an objective, absolute viewpoint, merely possible people and their surroundings are just as real as we and ours? Only if one identifies the objective or absolute viewpoint with a neutral standpoint outside of all possible worlds. But there is no such standpoint. The objective, absolute point of view is the view from the actual world, and it is part of our concept of reality that this should be so.[133]

From our viewpoint, all other viewpoints are just possible.

We could, however, separate the semantic analysis of 'actual' from the metaphysical thesis that actuality is a relation between a world and things in it. According to Stalnaker, one can accept one thesis and reject the other, just as one can accept an indexical analysis of personal pronouns and be a solipsist, or accept an indexical analysis of tenses, and yet believe that the past exists only in memory and the future only as anticipation.[134]

What is the picture of possible worlds we are left with? Stalnaker rejects Lewis's full-blown realism concerning possible worlds in favour of moderate realism. He rejects the metaphysical interpretation of the indexical analysis of actuality but accepts the indexical analysis as semantically correct. To him, possible worlds do not exist in the same way in which the actual world does - rather, he seems to treat them as a useful instrument.

> The concept of possible worlds that I am defending is not a metaphysical conception, although one application of the

---

[132]See Stalnaker, 1979, 228-229.

[133]At Stalnaker, 1984, 47. Even though the article 'Possible Worlds', Stalnaker, 1979, is basically the same as Chapter 3 of *Inquiry*, Stalnaker, 1984, some formulations are more precise in the later version.

[134]See Stalnaker, 1979, 229.

notion is to provide a framework for metaphysical theorizing. The concept is a formal or functional notion, like the notion of an individual presupposed by the semantics for extensional quantification theory.[135]

Stalnaker thus defends a formal notion of possible world. Just like the concept of an individual, he is happy to keep it as a primitive. The main attraction of Lewis's framework, as Stalnaker sees it, was its flexibility, and that flexibility he tries to preserve. By rejecting Lewis's extreme realism, Stalnaker avoids some problems inherent in that position, e.g., the theory of counterparts and the question of the domain of all possible worlds. Different kinds of possibility and necessity as well as different contexts shall give rise to various domains of different kinds of possible worlds, but there does not need to be any definitive or maximal domain from which all kinds of possible worlds are drawn.[136]

In his statement about the kind of possible worlds he wants to work with, Stalnaker manages to avoid at least some of the problems that plague both Lewis and Kripke. Indeed, if we accept both a 'possible world' and an 'individual' as primitive notions, it may seem that little more needs to be said on the subject. Stalnaker's framework seems remarkably free of metaphysical and ontological assumptions. The difference between Lewis's kind of ontological assumptions and those of Stalnaker is that Lewis asserts that existence of possible worlds without qualification. That makes his conception of possible worlds a metaphysical one. Stalnaker, on the other hand, acknowledges the usefulness of a possible-world approach to modality and avails himself of its principle advantages but avoids making an ontological commitment to possible worlds. By asserting that their existence is presupposed just for the sake of explanation, he puts it, so to say, in scare-quotes. This approach is not novel or limited to semantics. Hume[137] attacks the notion of causality but

---

[135]Stalnaker, 1984, 57.

[136]I defend this interpretation later. See also Stalnaker, 1981, 135, quoted in the next section.

[137]Hume, 1748.

admits its indispensability in everyday reasoning, and Adams proposes to treat modality by resorting to 'world-stories' without believing that world-stories really exist.[138] Regarding possible worlds, the variety of approaches to their existence is considerable, and a black and white classification (as in 'this author believes in possible world while that one does not') is grossly insufficient. We are not going to investigate the issues of approaches to existence of possibles in detail,[139] let us say only that the difference between Lewis's extreme realism about possible worlds and Stalnaker's approach is considerable, and that by treating possible worlds as tools within a theory of modality, Stalnaker does not seem to make a metaphysical commitment to their existence.

We shall now investigate some notions that are in Stalnaker's work connected with the concept of a possible world, and look for problems that could be tucked away there.

3.3.3   Proposition

At the outset, we noted that Stalnaker's main focus is on the explanation of intentional behaviour. It is important to add that for Stalnaker, a 'respectable' explanation is one given solely in non-intentional terms.[140] This is something we should keep in mind while going through the following sections. Right now, we shall turn our attention to analysing in greater detail the kind of possible worlds Stalnaker works with and the use he puts them to. We have already noted that, by rejecting Lewis's style of realism, Stalnaker escapes the problem of having to specify the domain of all possible worlds. In his framework, we have as many possible worlds as we need, no more, no less. Just how many worlds are needed, and of what kind, is determined by the context. We find an attractive statement of

---

[138] Adams, 1979.

[139] For a very nice overview of this discussion see Loux' extensive introduction to Loux (1979).

[140] See for example Stalnaker, 1984, 27: "I argued. . . that it is theoretically possible to solve the problem of intentionality - to give a naturalistic explanation of intentional mental states - without exploiting linguistic or semantic concepts."

this approach in his 'Indexical Belief', where Stalnaker says

> I doubt that it is plausible to believe that there is, independent of context, a well-defined domain of absolutely maximally specific possible states of the world... The alternative possibilities used to define propositions must be exclusive alternatives which are maximally specific, relative to the distinctions that might be made in the context at hand. But one can make sense of this requirement even if there is no ultimate set of possibilities relative to which any possible distinctions might be made. One might think of possible worlds as something like the elements of a partition of a space, rather than as the points of the space. The space might be partitioned differently in different contexts, and there might be no maximally fine partition. (This is only a rough analogy. And the space itself may also vary from context to context.)[141]

To better understand the relation between possible worlds and context, we first have to introduce the notion of proposition.

Stalnaker uses the notion of proposition to account for the content of representational mental states, e.g., beliefs, and also to express the impact of an utterance on the context and to capture the way an agent aims at changing something in the environment.[142] He uses a basically Kripkean notion of proposition, according to which a proposition is a function from possible worlds into truth-values.[143] This notion should be seen as set against the Russellian tradition, which holds that propositions are structured and reflect the structure of sentences that express them. Kripkean propositions are not structured in this way. Different sentences can express the same proposition, a particular sentence in different contexts can express different propositions, and, actually, a proposition need not even be expressed by linguistic means. Given a set of possible worlds, a proposition is determined by the subset of possible worlds in which

---

[141]Stalnaker, 1981, 135.

[142]See for example Stalnaker, 1998, 3: "...speech is action, and speech acts should be understood in terms of the way they are intended to affect the situation in which they are performed."

[143]Stalnaker, 1987, 2.

it holds. And this is where Stalnaker, but not Kripke, connects the notion of proposition with the notion of context - the context determines which possible worlds are relevant to determining the proposition. He says that

> If the alternative possibilities there are vary with the context, then so do the propositions which are, according to the conception of content I am sketching, just ways of distinguishing between the alternative possibilities. One can make sense of questions about identity and difference of the propositions expressed in different utterances or acts of thought only given a common context - a common set of possibilities that the propositions are understood to distinguish between. This yields a conception of proposition, which is less stable than, and very different from, the traditional conception, but it is, I think, more adequate to the phenomena of speech and thought.[144]

By making the relevant set of possible worlds dependent on the context, Stalnaker's propositions start behaving very differently from Kripke's. For example, in Stalnaker's framework, the notion of rigid designation becomes much weaker. Where in Kripke's framework a rigid designator denotes the same individual in all (metaphysically possible) possible worlds, in Stalnaker's framework a rigid designator designates the same individual in all the worlds of the context.

To sum up, according to Stalnaker, the proposition expressed by an utterance is a function over the set of possible worlds that are relevant to the context of an utterance – it divides those possible worlds into those where the proposition holds and those where it does not.

3.3.4   Context

We have now some idea of the role context plays in determining a proposition. It is time to turn our attention to the notion of context itself. Stalnaker offers a preliminary definition, saying:

---

[144]Stalnaker, 1981, 135.

> I propose to identify a context (at a particular point in a discourse) with the body of information that is presumed, at that point, to be common to the participants in the discourse.[145]

In general, there are two kinds of things that participants in a conversation assume to be shared. Firstly, they normally take it for granted that both speakers and hearers are aware that a conversation is taking place, and that all of them are aware of their surroundings.[146] Secondly, participants assume they share knowledge of some less immediate features that 'frame' the situation. For example, if they talk about American politics, they may assume that all of them know who is the president. Both kinds of assumptions contribute to the information which participants assume is shared. These are the presuppositions of the participants.

> We can represent the information that defines the context in which a speech act is taking place with a set of possible situations or possible worlds - the situations that are compatible with the information. This set, which I have called the context set will include all the situations among which the speakers intend to distinguish with their speech acts. The presumed common information - what is presupposed in the context - is what all these worlds have in common.[147]

A speech act of assertion takes place in all the possible worlds of the context set. In effect, an assertion is a proposal to add information to what is presupposed, that is, it is a move to eliminate from the context set those worlds that are incompatible with what is asserted.

Should we see presuppositions as objective features of the agents' interaction within a context or should we rather see them as necessarily tied to the perspective of one participant, usually the speaker? There is a degree of vacillation or perhaps just a lack of clarity in Stalnaker's work regarding this point but the latter seems to result

---

[145]Stalnaker, 1998, 5.
[146]Stalnaker, 1998, 5 and Stalnaker, 1978, 323.
[147]Stalnaker, 1998, 5.

in a more coherent reading. It is consistent with the repeated emphasis on information that is presumed to be shared (the presuming being done, one would suppose, by one agent at the time), as well as with the focus on the speaker, which we can witness in numerous quotations in this section. Even the passage where – as far as I could tell – Stalnaker defines the notion of presupposition for the first time supports the speaker-centred reading:

> A proposition is presupposed if the speaker is disposed to act as if he assumes or believes that the proposition is true, and as if he assumes or believes that his audience assumes or believes that it is true as well. Presuppositions are what is taken by the speaker to be the COMMON GROUND of the participants in the conversation, what is treated as the COMMON KNOWLEDGE or MUTUAL KNOWLEDGE.[148]

The structure of a speaker's presupposition can be represented by a Kripke model, in which the accessibility relation is serial, transitive, Euclidean, but not necessarily reflexive.[149] The requirement that it be transitive and Euclidian reflects the assmption that speaker presuppositions are transparent: speakers know what they are presupposing, so they that they are presupposing know they are presupposing P if they are, and that not if they do not. The requirement that the relation be serial reflects the assumption that the context-set is always non-empty, that there is always at least one possibility compatible with what is presupposed. The non-reflexivity is important here. It reflects the fact that some things the speaker presupposes may be false. This would happen either because an agent has a false belief or because she participates in some mutually recognised pretence. Moreover, it is not always, or perhaps not even usually, the case that participants in a conversation presuppose exactly the same things. An agent need not even believe that everything she presupposes is shared – it suffices if she pretends she does.[150]

---

[148]Stalnaker, 1978, 321, emphasis in the original.

[149]Stalnaker, 1998, 6.

[150]See the quotation immediately above as well as Stalnaker, 1978, 321, which is a direct continuation of the quotation above: "The propositions presupposed

Let us see now how Stalnaker puts the notions of context and proposition to work. His aim is to capture the effect the speaker's utterance has on the context, that is, on the beliefs of the audience. This he models with the help of a two-dimensional matrix called a propositional concept. We can best illustrate how it works using Stalnaker's own example:

> I said You are a fool to O'Leary. O'Leary IS a fool, so what I said was true, although O'Leary does not think so. Now Daniels, who is no fool and who knows it, was standing nearby, and he thought I was talking to him. So both O'Leary and Daniels thought I said something false: O'Leary understood what I said, but disagrees with me about the facts; Daniels, on the other hand, agrees with me about the fact (he knows that O'Leary is a fool), but misunderstood what I said. Just to fill out the example, let me add that O'Leary believes falsely that Daniels is a fool. Now compare the possible worlds i, j, and k. Here, i is the world as it is, the world we are in; j is the world that O'Leary thinks we are in; and k is the world Daniels thinks we are in.[151]

The following propositional concept corresponds to this situation:

	i	j	k
i	T	F	T
j	T	F	T
k	F	T	F

This is a two-dimensional matrix where on the rows we find either the proposition the speaker expressed or the proposition as a hearer understood it (this, of course depends on who is the speaker). For example, in the second row, j, we see the proposition as O'Leary understood it, evaluated in the world as he thinks it is, j, and in the worlds as the speaker and Daniels think they are (i and k). We construct a proposition as a function over the worlds that participants

---

in the intended sense need not really be common or mutual knowledge; the speaker need not even believe them."

[151]Stalnaker, 1978, 317-318.

think are relevant with respect to the situation at hand (this is of course as very simplified model). If we find the same proposition in two rows, we see that those two participants have understood an assertion as expressing the same proposition. In the columns, we find the worlds with respect to which we evaluate, that is, worlds in their role of a context (i.e., context worlds). Notice that in their role as context worlds, the world of the speaker and that of Daniels are the same (the verticals under i and k). That means that the speaker and Daniels assign the same truth-values to the proposition expressed as they each understand it.

On the diagonal, that is in squares ii, jj etc., we see whether the relevant participant himself thinks that the proposition as he or she understands it is true. In this particular propositional concept, we see in the square jj, that O'Leary, though disagreeing with the speaker about the facts, understands the utterance in the way the speaker intended. Daniels, on the other hand, does not understand the speaker's utterance the way it was intended. We can view the diagonal as expressing a proposition, so-called 'diagonal proposition', which has some interesting properties. A diagonal proposition is relevant to describing situation in which the speaker or the addressee has only a partial knowledge of facts that are relevant to determining what is said. For example,[152] if I get an undated postcard from San Francisco from my sister, saying, among other things, 'it is warm and sunny here today', I will not know exactly what it says (because I do not know when it was written), but I will know that it was written on a sunny day in San Francisco. The information I got from the postcard was the diagonal proposition of the propositional concept expressed by the writer, that is, my sister.

All this is well known. Yet, several more points need to be settled. Firstly, what determines which worlds occur in a propositional concept of a particular situation? This, in my view, is the place where the notion of context connects with Gricean maxims.[153] However simplified the propositional concept is, which possible worlds are

---

[152]This example is adapted from Stalnaker, 1999b, 13.

[153]Which are introduced in the famous Grice (1975).

relevant to modelling a particular situation is determined by a set of Gricean principles, mainly the principle of relevance. In building a propositional concept, we implicitly rely on the Gricean maxims, which we assume the participants have internalised.[154]

Secondly, we may find it surprising that in the quotation above, the world of the speaker is equated with 'the world as it is, the world we are in'. This seems to contradict the explicit set up of the notion of context, which is said to be the

> ...set of possible situations that are compatible with what is presupposed, or taken to be common ground, by the participants in the discourse.[155]

Here, context is seen as consisting of information that is presumed to be shared by conversation participants. We then model that information using propositions, which are a function over the possible worlds which the participants think are relevant. The content of an assertion depends on the context in two ways: firstly, the context is the object on which an assertion acts, and secondly, it provides the source of information needed to interpret an utterance.

> So both of the roles that contexts play require that they include a body of information: context-dependence means dependence on certain facts, but the facts must be available, or presumed to be available, to the participants in the conversation.[156]

Context consists of possible worlds that are thought by the participants to be compatible with the information they share. Stalnaker gives us no reason to suppose that one particular participant's context world is the actual one. In fact, the non-reflexivity of the model (which we mentioned on p. 132) rules that option out. Rather, if we look at the way Stalnaker describes context, it would seem that each participant thinks that his or her context set represents (a part of)

---

[154]We shall return to the importance of Grice's work to Stalnaker in section 4.4.5.

[155]Stalnaker, 1998, 7.

[156]Stalnaker, 1998, 5.

the actual world. This is not surprising once we go over the quotes
above and asks ourselves whether the worlds Stalnaker works with
are metaphysical or epistemic. It turns out that worlds consisting of
information presumed by the speaker to be shared are quite clearly
epistemic in nature. Nothing outside of epistemic alternatives can
get into them. It follows then that the content they describe is a
narrow, epistemic kind of content. But that does not seem to be the
kind of content Stalnaker had in mind when setting out to explain
intentional behaviour in non-intentional terms.

In order to get a fuller picture, we have to say more about
content, focusing on what Stalnaker wants to say about the non-
epistemic, broad component of it. And that is the topic of the
following section.

3.3.5   The Possible and The Actual

We saw that in analysing particular conversational exchanges, Stal-
naker uses epistemic possible worlds. Yet for any belief to be about
something other than just itself, it has to depend in some way on
the external world. Stalnaker addresses this problem by develop-
ing what he calls the 'information theoretic account of intentional
content'.[157] Using a different idiom, we can say that after devel-
oping a story about narrow content, Stalnaker gives us a theory of
broad content. To clarify his position, Stalnaker re-interprets the
best-known argument for broad content there is: Putnam's Twin
Earth thought experiment.[158] The Twin Earth thought experiment
- and a good many other thought-experiments inspired by it – was
designed to highlight the relation between the meaning of speaker's
utterances and the speaker's environment.

Even though it is well known, let us briefly describe Putnam's

---

[157]In the following, I discuss Stalnaker's views as he presents them in Stalnaker,
1993. The problem he deals with is by no means unique to his theory. It is a
version of the same problem that Block, 1986 is trying to address by making
narrow versus broad content distinction, and the same is treated in the discussion
about externalism versus internalism about belief.

[158]It comes from Putnam, 1975b.

example. Let us imagine a Twin Earth. The only difference between Earth and Twin Earth is that water on Earth is H_2O, while on the Twin Earth there is something that looks and behaves exactly like the Earth water but has a different chemical composition, which we shall call XYZ. On Twin Earth, XYZ plays exactly the same part H_2O plays on Earth. Let us also say that our scenario is happening before people knew about the chemical composition of water. Now, an actual, Earth speaker, O'Leary, has a twin on Twin Earth, a Twin O'Leary, who has all the same mental and physical properties as O'Leary does. Twin O'Leary should then have the same mental properties as O'Leary even when interacting with Twin water because there is nothing in his beliefs about that substance that could distinguish them from O'Leary's beliefs about Earth water. And yet, Putnam says, when O'Leary and his twin say something about water, they mean two different things. The meaning of their utterances differs depending on their environment. When O'Leary says 'There is water in the bathtub', he says something true on Earth, but would say something false on Twin Earth because there is no water there. Putnam argues that regardless of what Twin O'Leary thinks he means when using the word 'water', he cannot mean by it the same thing as O'Leary, who is an Earthling, because he, that is, Twin O'Leary, is not in position to refer to 'water'. 'Water' is a natural kind term, which in English inherently refers to H_2O. Twin O'Leary lacks the requisite causal connection with H_2O. That is why regardless of what Twin O'Leary thinks he means he cannot mean 'water' in the same sense in which O'Leary uses the term.

Putnam's point is about the source of semantic values. Sometimes,[159] the Twin Earth experiment was understood as showing that what one means is not a matter of mental states but of social conventions and causal connections. According to this interpretation, words can do their semantic work without us knowing what goes on. Thus while speakers' intentions and beliefs are inherently internal, their words depend for their meaning on external factors. This may conflict with our intuition that when a speaker is sincere

---

[159]See e.g., Fodor (1987).

and means what he says, then the speaker says what he believes.

Tyler Burge pointed out[160] that this is not the correct lesson to draw – according to him, Putnam shows that even belief and intention should be understood partly in terms external to the speaker. O'Leary's twin then not only says something different from O'Leary when he says 'There is water in the bathtub' but also acts on a different belief and a different intention. Burge's is a strongly externalist interpretation.

A theorist of an internalist inclination, on the other hand, tries to defend a position on which although what one sees and knows is partly a matter of external environment, what one thinks one sees and thinks one knows is a matter of the agent's internal state. In order to account for the role of the external environment, the internalist distinguishes between the way things seem and the way they are. The externalist insists that this is a remnant of Cartesian dualism. This is a serious objection because the internalist motivation is usually epistemological in the first place. The problem is this: if the internalist position entails a need for distinguishing between the way things are and the way they seem, his endeavour is jeopardised right at the beginning.[161]

Stalnaker is critical of the strongly externalist position. He feels that the claim that neither linguistic nor mental content is purely internal conflicts with certain common-sense intuitions. He says:

> Perhaps the externalist need not deny that what I say is what I think, when I am sincere, but it is hard to avoid the suspicion that if an externalist theory of speech and thought is right then we don't really know what we are either saying or thinking.[162]

Stalnaker's answer to the Twin Earth problems – the information theoretic account of intentional content – is an attempt to integrate the valid points of both the internalist and the externalist position. The idea is, basically, that

---

[160]In Burge, 1979.

[161]Stalnaker, 1993, 300-302. The sort of dualism he points to is exemplified for example in the work of Ned Block, Block, 1986.

[162]Stalnaker, 1993, 300.

> ...states of mind can carry information, when there exists a
> pattern of counterfactual dependencies between those states
> and corresponding states of the environment... Represen-
> tational states and systems carry misinformation as well as
> information in the strict sense, but according to the informa-
> tion theoretic picture, misrepresentation must be understood
> as a deviation from the norm. It is reasonable to assume that
> representational states are normally correct – that they are
> states that tend to represent things as they are.[163]

The dependencies between internal states and states of the environ-
ment have to be systematic because only then can we understand a
misrepresentation as a deviation from a norm. Normally, states of
mind tend to represent things as they are. An internal state rep-
resents the world as being such that P if under normal conditions
it would carry the information that P. An information carrying
state is a belief if it not only carries information but also monitors
the systematic dependency between an agent's environment and his
actions.

It may be very difficult, if not impossible, Stalnaker admits, to
give a non-circular account of normal conditions. For the time be-
ing, therefore, we have to make do with talking of a 'tendency to
carry such and such information'. Going back to the Twin Earth
example, we can say that what makes it the case that my beliefs
tend to depend on water is that water normally has certain observ-
able properties, and that it normally is the only kind of thing with
those properties around. These conditions do not obtain on Twin
Earth, and that is why there the same internal states fail to carry
information about Earth water.

The information theoretic story treats propositional content in
terms of causal regularities and counterfactual dependencies that
tend to hold under normal conditions. As to their particular form,
such regularities are contingent, and at least partly external. Inter-
nal facts carry information in virtue of how the world tends to affect
them. Therefore, if the world were different, the same internal states
would carry different information.

---

[163]Stalnaker, 1993, 302.

Using the information theoretic account, Stalnaker formulates his interpretation of the Twin Earth story: O'Leary's internal state as he steps into a bathtub is about water because, approximately, if what is in the bathtub were not water, he would not be in the cognitive state he is in. The internalist position, according to which O'Leary's beliefs are not about water but only about some stuff that looks like water (for example because he does not know that water is H_2O), would lead to the rather undesirable consequence of sceptical doubt.

Stalnaker's account predicts that O'Leary's beliefs are, under normal conditions, about water (or whatever water-like substance that normally occupies a certain role in his environment), and that his reference would fail only if normal conditions suddenly failed to obtain. This contrasts with Putnam's conclusions, according to which Twin O'Leary cannot have beliefs about water even though on the Twin Earth it is normal that the stuff that looks like water is XYZ.[164] This conclusion, Stalnaker would say, is undesirable because we might be well be in O'Leary's position, and on Putnam's view[165] our reference would then systematically fail. It would seem, therefore, that both the strongly internalist position and the position Putnam takes, lead to epistemologically unpalatable conclusions.

We have already noted that which conditions are normal is a contingent matter but the notion of normality itself is modal. In asking whether some condition is normal in a counterfactual situation, we ask either of two questions: 'Would it be normal in the actual world?', and 'Is it normal relative to that counterfactual situation itself?'

Once we make this distinction, we can get at the information theoretic answer to Twin Earth: Normally, O'Leary's internal state is about water (H_2O), because that is the main source of his information. If water (H_2O) suddenly changed to XYZ, O'Leary's reference to water would fail. On the other hand, if it were normal in the

---

[164]The same point could be made about the brain in the vat argument – it is normal for the envatted brains to be envatted, so why should they not be able to refer? Again, it is a crucial point of Putnam's story that they cannot refer.

[165]

actual world that water is XYZ, then XYZ would be what we would mean by 'water'.

We have to recognise that attribution of content is context-sensitive, and that what normal conditions are is relative to context as well. Attribution of content is essentially contrastive. When we ask whether the stuff in the bathtub is water, we presuppose that we have a good enough concept of water based on our normal conditions. To us, water is the stuff that is normally in the lakes, has certain observational properties, etc. And, as Stalnaker says,

> . . . attribution of content must be made relative to presumed facts about the background normal conditions. But any such can be called into question. The context of attribution can change, it changes what we say, but we don't have to change our minds.[166]

3.3.6   Content

We have seen that Stalnaker provides us with two accounts of content. Firstly, there is the content of assertions made in a particular context. That is what is described by propositional concepts. I have tried to show that this is, in fact, a narrow content.[167] Secondly, there is the notion of content derived from the information theoretic account. This notion of content is designed to capture the contribution of the environment, so it seems to be intended as an account of broad content. We should, however, keep in mind that the problem with which many theories of content struggle, and, indeed, the main task of a theory of content, is not just to describe the two kinds of content but, crucially, to explain the relation between them.[168]

Stalnaker explicitly addresses the notion of narrow content in several articles, arguing, broadly speaking, against the usefulness of

---

[166]Stalnaker, 1993, 309.

[167]The term and concept of narrow content is introduced in Block (1986).

[168]We can express this in various idioms, e.g., in terms of a relation between internalist and externalist intuitions, coherentist versus metaphysical theories of truth, internal versus external anchors, but the core of the problem remains the same.

the notion.[169] The explicit aim of his project with respect to content is to develop a notion that bridges the gap between the narrow and the broad notion. If, however, what I argued for in previous sections is correct, then it would seem that what Stalnaker says he wants to do and what he actually does are two different things. That the actual work does not quite live up to the explicit project happens often enough and it need not trouble us *per se*. It would be troubling, however, if we found out that the two notions of content Stalnaker develops clash.

In order to clarify the relation between the notions of narrow and broad content as Stalnaker describes them, we shall have to explain his use of diagonalisation. This brings us to the *locus classicus*, the place where an analysis of 'O'Leary believes that Hesperus is Mars' is proposed.[170]

In brief, the example runs as follows: O'Leary believes that Hesperus is Mars. He chooses to express his belief by saying 'Hesperus is Mars'. Stalnaker analyses this assertion as follows: Let us suppose that Kripke convinced us that proper names are rigid designators, that is, that they refer to the same object in all possible worlds. A proposition expressing identity between two rigid designators, here the proper names 'Hesperus' and 'Mars', is then either necessarily true or necessarily false. Indeed, this example targets two points: one about the behaviour of proper names, the other about believing in necessary propositions.

In constructing a propositional concept of the assertion 'Hesperus is Mars', we take into account the world O'Leary believes to be in. Because O'Leary asserts what he believes, in his world the assertion is true. And because his assertion involves the identity of two rigid designators, it is necessary. In the actual world – as anyone who ever studied identity statements knows– this assertion is false, and necessarily so. The propositional concept then looks as follows:

---

[169]For example in Stalnaker, 1989, and Stalnaker, 1990.
[170]Stalnaker, 1987, 179.

	i	j
i	F	F
j	T	T

> In j, O'Leary's world, 'Hesperus' designates Mars, while in i, the actual world, it designates Venus.

There is a problem, Stalnaker notes,[171] with ascribing a belief in a necessarily false proposition to an agent. Intuitively, one should think that speakers assert things they believe to be true. That is why in some cases, Stalnaker says, the content of an assertion is best described by a diagonal proposition (which we can read off squares ii, jj, etc.) derived from a propositional concept. The content of assertion is to be identified with the diagonal proposition in case the 'standard' interpretation would violate Gricean maxims, that is, when we would have the speaker assert something extremely implausible. In identifying the meaning of O'Leary's utterance 'Hesperus is Mars' with the diagonal proposition of the propositional concept, we emphasise the fact that he made his utterance in ignorance of certain facts that were relevant to what he meant to say. In doing this, we avoid the need to saddle him with asserting a necessarily false proposition.

This example has a number of puzzling features. The first concerns the perspective from which it is constructed. In a previous example of a propositional concept, in the case of O'Leary being called a fool, the situation was described from the viewpoint of a speaker. A propositional concept is intended to capture the impact that a speaker's assertion has on the belief alternatives of his audience. However, in this case, there is no audience. O'Leary is not imparting his views to anyone in particular. The situation involves just him and the actual world.

I think the preferred way of interpreting this propositional concept involves the assumption that it describes a situation from an outside point of view. In this case, the speaker is not O'Leary but someone describing the content of O'Leary's beliefs. That person

---

[171]Stalnaker, 1987, 179.

then also makes assumptions about what the actual world is. So even though this example was not intended as such, it is in fact a case of attitude ascription.

Further on, as I just mentioned, one of the possible worlds in the propositional concept above is identified as 'the actual world'.[172] Given the way the notion of propositional concept is set up, this seems rather surprising. Generally speaking, a propositional concept is supposed to represent context, that is a body of information or a

> ... set of possible situations that are compatible with what is presupposed, or taken to be common ground, by the participants in the discourse.[173]

If this is the case, then how can a particular viewpoint be identified with 'the actual world'? What is the role of the actual world in a propositional concept?

It may be tempting to dismiss the use of 'actual world' in this particular example involving Hesperus and Mars as a slip of the pen. We could do that if this were the only place where Stalnaker invokes the 'actual world' in connection with a propositional concept. But we have already noted above, when first introducing the notion of propositional concept, that the viewpoint of the speaker was equated with the actual world there as well. In that example – where the speaker calls O'Leary a fool – the perspective of the speaker is identified with 'the world as it is, the world we are in'.[174] In another example, when analysing a stipulation that 'Julius' is the name for the person who invented the zip,[175] Stalnaker equates the actual world with the world where the stipulation takes place. A similar thing happens in the analysis of the statement 'Sherlock Holmes does not exist'.[176] Given this evidence, we cannot dismiss the occurrence of 'the actual world' in propositional concepts as a slip of the pen. It figures in most examples of propositional concepts that Stalnaker gives in his work.

---

[172]Stalnaker, 1987, 184.
[173]Stalnaker, 1998, 7.
[174]Stalnaker, 1978, 317.
[175]Stalnaker, 1999b, 15.
[176]Stalnaker, 1978, 330.

In looking for an answer to this puzzle, we should start by summing up some relevant characteristics of Stalnaker's possible worlds. The possible worlds of a propositional concept are said to represent the context, that is the worlds that individual participants think they could be in. As we pointed out earlier, the notion of context is closely connected with the notion of presupposition and the notion of belief.

> In general, to understand the content of a person's belief, ask what the world would be like if the belief were correct. What is the world like, according to the person's conception of the way the world is? If we can give a coherent account, in our own terms, of a way, or a set of ways, that things might have been which seems intuitively to represent correctly the person's conception of the world and his place in it, then we will have explained his beliefs and attitudes to propositions - objects which conform to the received doctrines.[177]

We can thus assume that each participant thinks that her context set includes a representation of part of the actual world in the sense of a broad content. But it is important to keep in mind that, as Stalnaker repeats in a number of places,

> ...according to the conception of content I am presupposing,...a person thinking [a] thought will be an inhabitant in each of the possible situations compatible with his knowledge.[178]

This captures the insight that the agents themselves cannot be quite sure which of the possible worlds compatible with their knowledge is, or is closest to, *the* actual world in an externalist sense. In order to understand Stalnaker's use of the notion of actual world within the context of a propositional concept, we have to look again at his reasons for accepting a possible world framework in the first place.

When analysing Lewis's claims about possible worlds, Stalnaker rejected the metaphysical interpretation of indexicality of 'actual'.

---

[177]Stalnaker, 1981, 135.
[178]Stalnaker, 1981, 142.

He said that "the objective, absolute point of view is the view from
the actual world, and it is part of our concept of reality that this
should be so."[179] Therefore, we might add, it follows from our con-
cept of reality that we identify our particular point of view with
'the actual world'. The speaker takes his or her point of view to
be the actual world, even though he or she may entertain doubts,
be uncertain, or lack information regarding various features of it.
Possible worlds are introduced to model our attitudes to the world
that surrounds us.

Given the strongly epistemic flavour of the motivation and use
of possible worlds in Stalnaker's work, it is only fair to ask whether
his position can be described as realist at all. Stalnaker asks himself
the same question:

> Is the form of realism about possible worlds that I want to
> defend really realism? It is in the sense that it claims that
> the concept of a possible world is a basic concept in a true ac-
> count of the way we represent the world in our propositional
> acts and attitudes.[180]

We can now conclude that the 'actual world' that figures in Stal-
naker's propositional concepts is an epistemic concept that captures
the perspective of the speaker.

It remains now to see how narrow content – for we have seen
now that the content described in propositional concepts is, indeed,
narrow – connects with broad content described in the information
theoretic account of intentionality.

3.3.7   Narrow and Broad Content

Let us recall that the problem with an account of narrow content is
that unless it is complemented with an account of broad content, it
is hard to explain why agents' beliefs are about the external world
at all. We shall now look at how Stalnaker handles this problem.

---

[179]Stalnaker, 1984, 47.
[180]Stalnaker, 1979, 234.

In Stalnaker's information theoretic account (as we described it in section 3.3.5), the content of the agents' beliefs can be about the actual world (in the externalist, broad sense) because that is the environment agents are normally in. On the information theoretic account, the actual world is the main source and reference point of the information an agent possesses. In fact, whatever is the main source of information for an agent is that agent's actual world.

If, for example, an object x normally causes an agent to think that x is present, then that agent's beliefs are about x, and x is the source of information about x. The information theoretic story is about general norms: normally, beliefs about x are caused by x.

In particular situations, such as those we find described by propositional concepts, the information theoretic account cannot help us decide whether a particular piece of information is true. When building a propositional concept, we assume that the worlds participants think they are in can vary in details, even in details that are relevant to the situation at hand. Usually, all conversation participants are in the same physical surroundings but their perspectives, their perceptions of what the 'actual world' is, may vary.

The information theoretic account tells us why people's conversations, though happening on the level of epistemic possible worlds, are about the external world. It cannot help us decide who is right in a particular case. It does, however, explain why people's conversations are not just about beliefs – it explains why beliefs are about the world that surrounds us. This is also the point where communication connects with action: if we see speech as a kind of action, the information theoretic story explains how it can have an impact not only on the beliefs of the audience but also produce a desired effect in the actual world.

As it stands, the information theoretic story works on a different level and is much more global than the analysis of assertions. The final picture we are left with, accounts for narrow content of speaker's assertions in particular situations, and for broad content in a general framework. It has been proving very difficult to establish a link

between narrow and broad content.[181]

Stalnaker does not seem troubled by the narrow versus broad content problems. His solution to the Twin Earth thought experiment aims at bridging the gap between the content conceived of as internal to the speaker (marrow content) and content as an inherently public phenomenon (broad content). We may feel that regardless of Stalnaker's own perception of what he suggests, the proposal concerning the Twin Earth is internalist from the word go. Stalnaker is not likely to see it that way. This is because he never explicitly acknowledged that the possible worlds in the propositional concepts are epistemic in nature, and that the content they describe is thereby necessarily narrow. Especially when discussing necessary propositions and semantics of rigid designators, Stalnaker seems to think he still works within a Kripkean framework, as if not fully accepting the ramifications of the differences between his and Kripke's framework. That is somewhat regrettable.

My analysis of the connection between narrow and broad content, however well supported by a previous study of the key notions, is therefore necessarily largely speculative.

3.3.8   The Resulting Picture

Over the course of this subchapter, we have gathered a number of somewhat puzzling observations. On its own, each of them is just a little crack in the surface of the standard interpretation of Stalnaker's work. By the standard interpretation I mean the approach that sees Stalnaker as working with a notion of context as a set of participants' presuppositions, and reads the possible worlds he works with as rather akin to Kripke's possible worlds. On this interpretation, Stalnaker's attitude to context is often likened to Kaplan's. This 'standard' reading can be found, for example, in the works, of Ulrike Haas (Haas-Spohn, 1994, Chapter 2), Craige Roberts(Roberts, 1996), Massimo Poesio (Poesio and Traum, 1997), Alessandro Capone (Capone, 2003), and Ben Caplan (Caplan, 2004).

---

[181]See for example the analysis in Lepore and Fodor (1992).

A careful examination of the central notions of Stalnaker's approach to semantics has led me to a different reading. Its presentation shall follow the cues of questions that were left unanswered in the previous sections.

When characterising Stalnaker's motivation (section 3.3.1), we pointed out that it is likely that Stalnaker's main problem will be to find and describe the right balance between internalist and externalist concerns. For Stalnaker, this is crucial because an investigation of representational mental states and their connections to the surrounding world is at the very core of his enterprise. Later, in section 3.3.5, we saw that as a general theory of content, he endorses a version of moderate externalism. We ran into problems when we tried to connect his overt agenda about content (which we found in Stalnaker, 1990 and Stalnaker, 1993) with some implications of his analysis of conversation. In analysing his notion of a possible world (section 3.3.2), we concluded that Stalnaker adopts a very weak version of realism. Actually, he uses possible worlds to model attitudes of conversation participants. Stalnaker's possible worlds are basically epistemic in character. In a sense, for Stalnaker, the actual world is 'I and my surroundings', but it is a different sense than Lewis tried to argue for.

In the last two sections, I outlined a somewhat speculative interpretation of the connection between Stalnaker's notions of narrow and broad content. Having shown that the content characterised in propositional concepts is narrow, and the content derived from the information theoretic theory is broad, I tried to see how they fit together. I conjecture that the broad, information theoretic account, is intended to underpin narrow content. The theory of how broad content is derived is too general to be of practical use in analysing particular situations. This is not just because it is only roughly sketched: it is inherent in its very set-up, and rightly so since it may well be the case that, in principle, external circumstances fail to fully determine the mental state of an agent. However, we can make assumptions about some things that agents normally believe in given external circumstances. This is why the epistemic content of agents' beliefs (and, consequently, their utterances) can be seen

as linked, in an underdetermined and general way, to broad content and the external world.

Why did I not draw on other people's work in presenting my interpretation of Stalnaker? Why was I able to find only one person whose views on Stalnaker are somewhat similar to mine?[182] I suppose it is because Stalnaker himself does not make much explicit connections between the various parts of his work. What we have are articles and passages that present the general program (e.g., his views on narrow content), and then passages that deal with particular problems using a different framework (e.g., propositional concepts). Most people who use Stalnaker's work focus on particular applications of his two-dimensional semantics to descriptive semantics, mainly in natural language semantics. My goal here, however, was to trace Stalnaker's use of possible worlds and his notion of content throughout most, if not all, of his writings. In absence of explicit connections between various parts of Stalnaker's writing, I had to speculate to a degree. But it has been my aim to try and figure out what Stalnaker would have said had he focused on the questions that I am dealing with.

3.4   Conclusion

In the course of this chapter, we analysed the possible world frameworks of Lewis, Kripke, and Stalnaker, focusing on the aspects of their proposals that have consequences for the treatment of proper names. We treated each proposal separately because that way we could best uncover the internal coherence of each, and that, in turn, allowed us to describe some less obvious features of these proposals. Our aim was to see how the particular details of setting up a possible world framework influence the way proper names and their content are treated.

We saw that a strongly realist approach to possible worlds – such as we encountered in Lewis's work – necessitates the adoption

---

[182]I have in mind a paper by Peter Alward (Alward, 2004) , which draws on his Ph.D. thesis.

of worldbound individuals, which, in turn, has consequences for the applicability of rigid designation, making it trivial and of little if any use. Thus while we concluded that Lewis's framework is very general, has amazingly much internal coherence, and is, at first sight, free of essentialist commitments, its trademark feature, strong realism about possible worlds, has pervasive consequences that make its applicability to the study of modal properties of proper names rather limited. In the next chapter, in sections 4.1.3 through to 4.1.5, we shall see that the strongly realist approach to possible worlds, coupled with absence of any in-built essentialist features, has pervasive and rather undesirable repercussions for a foundational semantics based on this approach.

In our study of Kripke's proposal, we concluded that the main disadvantages of the Lewisian conception were absent. Kripke's framework allows for transworld individuals, and makes – not surprisingly –good sense of rigid designation. We saw that in order to make the minimal rigid designation work, one has to presuppose an essentialist version of haecceitism. That, *per se*, we did not find objectionable. What gave us a cause for concern was the scope and scale of essentialist theses and principles regarding particular natural kinds (e.g., the kind 'water' or 'gold'), the adoption of which was not necessitated by the rigid designation as such. We argued that such metaphysical theses cannot be derived from an analysis of natural language. Furthermore, metaphysical necessity, closely connected with Kripke's style of essentialism, turned out to be a very important feature in determining the domain of Kripkean possible worlds. Altogether, we found the part of Kripke's framework that relies on an adoption of scientific realism objectionable. Moreover, we shall see (in section 4.2) that the descriptive semantics Kripke proposes for treatment of individuals causes problems on the level of foundational semantics.

Our reconstruction of Stalnaker's proposal was to some degree speculative. While we tried to stay within the spirit of his proposal, sometimes we had to hypothesise what his answer to a particular problem would be. We saw that Stalnaker works with epistemic possible worlds but also has a story about the connection between

the epistemic content and broad content. All and all, of the three proposals, we found his to be the best suited to the study of the functioning of proper names in a natural language. This we shall be able to confirm later, when reconstructing his proposal for the foundational semantics of names in section 4.3. Our own proposal for the descriptive semantics of proper names (chapter 5), too, shall bear marks of Stalnaker's influence.

We have now the opportunity to make some remarks on the status of the semantics of modal statements. We saw both in Lewis's and in Kripke's case that building a framework largely in order to capture modal intuitions is a risky enterprise. In Lewis's case, his desire to say something about the relative likelihood of various counterfactual situations (and, perhaps, an independent belief in the existence of possible worlds) lead to a development of a framework which seems to have little to contribute to natural language semantics. In Kripke's case, his modal intuitions were incorporated into a system that was designed to account for natural language phenomena. This, however, resulted in a potential clash between our linguistic and modal intuitions, which we regard as an undesirable consequence. Thus is seems that while a possible-world framework has to allow for the expression of some of our modal intuitions, it should be general enough to accommodate the differences between speakers. Idiosyncratic modal judgements should not be built into the framework if it is to play a more general role in the semantics of natural language. A possible-world framework should primarily be a suitable medium for capturing linguistic intuitions (for example, rigid designation of proper names), and the influence of modal judgements should be kept at a minimal level.

Throughout this chapter we kept our focus on the connection between the possible world set up, its essentialist preconditions, the notion of an individual, and the kind of necessity that ties these notions together. In the following chapter, we shall continue our investigation of necessity, bring into focus the notion of a proposition, and see how various frameworks deal with the questions raised by the extensional character of propositions, that is mostly the problems of logical omniscience and necessary identity statements.

Chapter 4
Foundational Semantics

The goal of this chapter is to examine some key notions of theories of reference of proper names, mainly the notion of proposition, from the perspective of foundational semantics. As we said earlier, in section 1.4, the task of foundational semantics is to look at what it is about speakers' behaviour that endows different kinds of expressions with the semantic value they have, or rather, the semantic values the descriptive semantics predicts they should have. An important part of the foundational undertaking is to look at how the speaker communicates her intentional states (e.g., beliefs) in her utterances so that these have the desired effect on the speaker. This is why we shall focus on what the frameworks we have been investigating, that is Kripke's, Stalnaker's and Lewis's framework, have to say about the content of utterances. The relation between what one can assume that the speaker intends to communicate and what a theory predicts she actually manages to get across is one of the main points we shall address.

As we had already indicated, any account of modal and descriptive semantics of a particular kind of expression can be successful only if it makes plausible predictions about the prerequisites the speaker has to meet in order to use that kind of expression successfully. By focusing on the notion of proposition, we shall, in a natural way, extend our investigation of descriptive and modal semantics of proper names. This is because, as we shall see, the lessons of pre-

ceding chapters have rather immediate consequences for an account of propositional content. An account of content, while being a part of foundational semantics, is especially closely connected to the account a theory puts forward to deal with modal statements. This we shall see at many points in this chapter.

Our investigation is going to be epistemological – we shall focus at the epistemological consequences that setting up a possible-world framework in a particular way has with respect to the notion of proposition. Where we encounter particularly important problems, we shall investigate various suggestions that were put forward to amend the situation. This will especially be true of the Pierre puzzle. As before, we shall proceed author by author. This will allow us to fully appreciate the internal coherence of various approaches to semantics of proper names, and see which of the approaches seems best suited to be a basis of our own proposal.

4.1   Believing in Propositions

Some of the theories we have been analysing so far – as well as many we have not mentioned – take the notion of proposition to be central to describing the content of utterances. However, the particular notion of proposition we have been working with seems to have some features that might make it difficult for actual natural language users to use propositions in the way they are supposed to, that is, to express their intentions, beliefs, and desires. We shall now look at those problems, and see to what extent the frameworks we have described in the preceding chapter suffer from some counterintuitive consequences of their notion of proposition.

So far, we have assumed a notion of proposition according to which a proposition is defined as a function from possible worlds into truth-values. This is basically the approach that has been pioneered by Kripke. To fully understand this definition and its implications, one should look at the concepts on which it is based. Firstly, one should take into consideration the kind of possible-world framework this notion is applied to. We have examined in detail Lewis's,

Kripke's and Stalnaker's possible-world frameworks in the chapter on modal semantics. The results of that chapter shall be of use to us now.

Secondly, one should appreciate the extent to which the notion of a truth-value that is used in defining a proposition is, or is not, important. Basically, the most noteworthy thing about truth-values here is that there are only two of them – true and false. We could reformulate the definition of proposition leaving out all reference to the notion of truth. We would then view a proposition as a subset of some domain of possible worlds. The role of truth-values is simply to distinguish those possible worlds that are a part of the selected subset from those that are not. In other words, a proposition is fully defined, relative to a domain of possible worlds, by a subset consisting of those possible worlds for which it has the value true. Therefore, if two propositions are defined for the same arguments, and take the same value for each argument, they are identical. Propositions are fully extensional functions.

This has some far-reaching consequences. Propositions are often used to characterise the objects of intentional mental states, for example beliefs. But it follows from our definition of proposition that an agent who believes a particular proposition believes by the same token all propositions that are identical with, that is, necessarily equivalent to, the believed one. This is called the problem of logical omniscience. It is especially pressing when we turn to necessarily true propositions. A necessarily true proposition is true in every possible world under consideration. Therefore, all necessarily true propositions (in a given domain) are necessarily equivalent. Hence, there is really only one necessarily true proposition, the tautology. By believing in one necessarily true proposition, the speaker is predicted to believe all of them. That is a very counterintuitive result.

One area of inquiry where the problem of logical omniscience has vast consequences is mathematics. The way the notion of proposition functions, implies that if the statements of mathematics are necessary – as many philosophers would indeed claim – then the whole of mathematics consists of one true proposition, which can be expressed in many different ways. I will not deal with mathemat-

ics in this chapter or anywhere else. An analysis of the particular problems of mathematics within a particular semantic system is a chapter of its own, and it does not fall within the focus of the present enterprise.[1]

4.1.1   Do We Believe in Necessary Propositions?

The problem of logical omniscience is very simple and convincing. Basically, if propositions are indeed functions from possible worlds into truth-values, how can they be objects of intentional states or even just their adequate characterisations? What strategies can be adopted to deal with this problem?

One possible reaction would be to reject this coarse-grained notion of proposition – according to which propositions do not have an internal structure – and endorse instead one more finely grained. There are many philosophers who have taken this route, Hartry Field[2] being currently perhaps the most prominent advocate of the fine-grained notion.[3] The basic idea of the fine-grained approach is that some problems can be avoided is we assume that a proposition has a structure that mirrors the structure of the sentence that is used to express it. A proposition is then seen as an ordered set of things, properties and relations. This approach has its advantages, breeds its own problems, and as to its practicality the jury is still out.[4] The motivation behind this notion of proposition is that by

---

[1]The issues sketched here have generated an enormous amount of discussion whose records would fill libraries. I shall only focus on those features of the discussion that have a direct bearing on foundational semantics of proper names, that is, I shall be looking at the links between a descriptive account of proper names, the ways this influences the set-up of a possible-world framework, and the consequences this has for the notion of a proposition, with an emphasis of the foundational semantic issues.

[2]See for example Field, 1977.

[3]It is, of course, quite a bit of simplification to talk about *the* fine-grained notion of proposition. There are good many versions around. Our purpose here is, however, only to sketch the idea on which this alternative approach is based.

[4]The literature dealing with fine-grained approach to content and structured propositions is very extensive. Among the important works are Fodor and LePore, 1992, Cresswell, 1985, King, 1996.

enabling us to distinguish more finely what it is that agents believe, it should provide us with more plausible objects of intentional states. However, its main attraction is also the cause of its main problem: where the coarse-grained approach, such as the one where we individuate propositions by truth-values, sees too many equivalencies between propositions, the fine-grained approach may see too few. As a consequence, it may be difficult for the advocates of this notion to explain why different sentences can play the same role in agents' reasoning.

However, so far we have only dealt with frameworks that use a coarse-grained notion of proposition, and we shall continue this trend. We shall now look at various defensive and evasive strategies that have been proposed to deal with the problem of logical omniscience, and with the problem of unwanted identities between propositions, within that approach. We shall look again at the possible-world frameworks we examined before, and see whether and to what extent they are actually vulnerable to the problems outlined above.

4.1.2   Why Put Up With Logical Omniscience?

In Stalnaker's recent work, we find an attractive classification of approaches to the problem of logical omniscience.[5] I shall briefly introduce it here because keeping it in mind will help us see what various authors are doing, and shed light on the basic strategies they use to deal with the problem of logical omniscience.

There are basically two ways, Stalnaker says, to reconcile the fact that people do not believe all logical consequences of their beliefs with a theory that predicts that they do.

On the one hand, one may interpret the notion of belief in one's logic as being about belief in the ordinary sense of the word, but restrict the domain of its application to idealised believers. These believers would have unlimited memory capacity and infinite computational ability and speed. Ordinary people cannot think of ev-

---

[5]The presentation of this section follows Stalnaker, 1999c, 242-245.

erything – these idealised believers could.

On the other hand, one may leave the domain of application unrestricted, thus including ordinary agents such as you and I, but reinterpret the concept of belief to which one's logic is intended to apply. One would then assume a difference between an ordinary belief and a belief in some technical sense. One could, for example, distinguish belief in the ordinary sense from *implicit belief*, where one's implicit beliefs would include all deductive consequences of one's beliefs. Crucially, one would not have to assume that these implicit beliefs are accessible to the agent to whom they are ascribed. Being logically omniscient with respect to such implicit beliefs would be easy – in fact, even the least reflective agent could not help being logically omniscient in this sense.

Both of these idealisations begin by conceding that agents cannot believe, in the ordinary sense of 'believe', all logical consequences of their beliefs. But why should one try to idealise the situation in the first place? Why not try to develop a logic that describes beliefs in the ordinary sense as held by ordinary agents? There are, basically, four different motivations for idealising belief in either of the ways sketched above.

Firstly, one may idealise in order to get at the underlying mechanisms. Complicated behaviour of a system may be explained by an interaction of various mechanisms that can best be understood by looking at how they work in isolation. A theory may focus on one component, and view the action of other components as external interfering factors. The assumption of logical omniscience is sometimes seen, at least implicitly, as an idealisation motivated in this way. Failures to believe all the consequences of one's beliefs are seen as resulting from a kind of cognitive friction that impedes the natural process of drawing consequences. This, Stalnaker says, is an attractive picture, but we shall see that it rests on an implausible conception of what belief and knowledge are.[6]

A second reason for idealisation can be simplification. It may be that some features of a system that complicate its accurate de-

---

[6]Stalnaker, 1999c, 244.

scription can be, for some purposes and in some contexts, neglected. It has been suggested that the assumption of logical omniscience in normal epistemic logics is of this kind. Robert Moore, for example, says that logics that imply logical omniscience

> ... represent idealizations that are reasonable approximations to the truth for many purposes. While no rational agent's knowledge is closed under logical consequence, outside of mathematics there seem to be few cases where this significantly affects an agent's behavior.[7]

Yet this, Stalnaker claims, is an underestimate of the extent of the distortion. Any context where an agent reasons is a context that is distorted by the assumption of logical omniscience since reasoning (at least deductive reasoning) is an activity that deductively omniscient agents would have no use for. To such idealised agents, deliberation is unnecessary. In fact, any kind of processing or computational activity is unintelligible as an activity of a deductively omniscient agent. It is hard to see the adoption of deductive omniscience as a harmless simplification if the logic of knowledge that adopts it cannot say anything about such essential activities as reasoning and deliberation.

A third kind of justification for idealisation is normative: whatever the actual divergence of real agents from logical omniscience, is it not something that rational agents should strive to approximate? This suggestion seems, at least *prima facie*, plausible but a number of philosophers have suggested that rational agents may have reasons to avoid accepting all the consequences of their beliefs. For example Gilbert Harman says that 'many trivial things are implied by one's view which it would be worse than pointless to add to what one believes.' He proposes a principle of reasoning he calls 'clutter avoidance': 'One should not clutter one's mind with trivialities.'[8]

Stalnaker says that even if deductive knowledge were a normative ideal, that would not be a reason to build it into one's theory of knowledge.

---

[7]Moore, 1988, 363.
[8]Harman, 1986, 12.

The fourth reason for idealising is more pessimistic. Perhaps it is just that at the moment we cannot do any better than to idealise either the believer or his beliefs. Stalnaker concludes that this reason comes closest to the truth. We shall now look at various possible-world frameworks and see to what extent the problem of logical omniscience arises within them. We shall also look at how they try to deal with the problem.

4.1.3 Propositions and Worldbound Individuals

We shall start by looking at the way a strongly realist possible-world framework deals with the problem of logical omniscience. In doing so, we shall draw strongly on our previous study of Lewis's work (subchapter 3.1). The following investigation is not intended to show what Lewis's own approach to foundational semantics is. It is meant to show what the consequences of a strongly realist possible-world framework are with respect to foundational semantics. In the following section (section 4.1.4) we shall briefly outline Lewis's actual position to foundational semantics. For the time being, let us just note that Lewis developed his possible-world framework to suit the analysis of counterfactual statements, not in order to capture the content of people's intentional states. His own analysis of objects of belief does not rely on the notion of proposition under investigation here. We chose his framework in order to investigate what happens with the problem of logical omniscience in a strongly realist possible-world framework, and of that Lewis certainly is an advocate.

Two features play a prominent role in distinguishing a strongly realist possible-world framework from its less realist brethren. Firstly, there is the worldboundedness of individuals within it. Because a non-actual possible world exists in the same sense in which the actual world does, an individual, say Samuel, can only exist in one world. Surely, there are individuals in other possible worlds that resemble Samuel arbitrarily closely but they are, strictly speaking, distinct from him. This claim is at the core of the theory of counterparts.

The other important feature of a strongly realist approach to possible worlds is that it is, at least in its general form, profoundly

non-essentialist. There are no properties, save formal ones, that an individual and his counterpart in another possible world have to share.[9] We dealt with these claims in section 3.1.2. Let us now look at the consequences these two features of a strongly realist framework - the worldboundedness of individuals and absence of essential properties – have for the possible-world-based notion of proposition.

We shall look at three kinds of sentences, those expressing identity ($a = b$), those expressing predication (Pa) and those expressing general predication (e.g., 'all a's are b's.'). Sentences expressing identity between singular referring terms, such as 'Hesperus is Phosphorus' are – if we take those terms to be rigid designators – usually taken to express necessary propositions. If an identity statement is true, it expresses a necessarily true proposition. As we just noted, in a strongly realist framework, individuals are worldbound. This follows from tenet (b) of the theory of counterparts (see section 3.1.2): Nothing is in two worlds. Both Hesperus and Phosphorus therefore exist only in one world, in this case the actual one. The statement expresses a proposition that is true in every possible world where the referent occurring in it exists. Here, however, the referent exists always only in one world. As a consequence, the necessity connected with an identity statement is always limited to one world. This is somewhat strange but not really counterintuitive.

What is surprising, however, is that something very similar happens with predication in sentences such as 'John is tall'. Here the individual, John, exists only in one world, in our case the actual world, and is described as having some property, e.g., 'being tall' in that world. However, if John is indeed tall in the actual world then per definition he is tall in every world in which he exists. Lo and behold, 'being tall' just turned out to be one of John's essential properties. If fact, every property that can be truly predicated of an individual in a given world is essential for that individual. These examples also illustrate that the notion of rigid designation for sin-

---

[9]An individual is also a counterpart of itself but for the time being I am interested only in those counterparts that are not in the same world as the individual they are a counterpart of.

gular referring terms finds little use in a framework of worldbound
individuals.

We can generalise now and say that in a strongly realist frame-
work that has the features outlined at the beginning of this section,
necessity collapses for all propositions expressing predication to and
identity of individuals. That is, not only is it the case that all iden-
tity between individuals, if true, is necessarily true, but the same
holds of all properties of individuals. Having a red tie on is essential
just as much as being human is. This is a proximate consequence
of having worldbound individuals. We have lost, with respect to
individuals, the distinction between necessary and contingent prop-
erties. The 'traditional' problem with necessary statements is that
it is predicted that in virtue of believing in one necessary statement
of the form Pa, one believes in all necessary statements of that form.
In this particular case, because all statements involving individuals
are necessary (in a particular way given by the world-boundedness of
individuals) but only true in one world, the whole problem of logical
omniscience becomes very hard to even comprehend.

Let us now turn our attention to general predication, to sentences
like 'All cows are mammals'. Cows may exist in various possible
worlds and, depending on the way the counterpart relation is set
up, even though they may be mammals in some, they may turn out
to be automata or look just like eggs in others. While the strongly
realist framework delivers masses of essential properties to individ-
uals, it assigns none to classes or kinds of individuals. In fact, given
the variety of shapes and forms the counterparts of some actual-
world kind may assume, it is hard to see what holds members of
a kind together across possible worlds save the fact that they are
counterparts of each other.

One could object that this strict reading of worldboundedness
is uncharitable. Let us take this objection seriously and concede
that though Lewis consistently defends the notion of worldbound
individuals, an individual and its counterpart are representations
of each other. Let us assume that even though strictly speaking I
cannot hold a counterfactual belief about an actual-world individual,
for example Mark Twain, because such belief involves Mark Twain's

counterpart, having this counterfactual belief is almost the same, or close enough to having a belief about the actual-world Mark Twain. Once we made this concession, we can say that a 'rigid designator', e.g., the proper name 'Mark Twain' will refer to all of Mark Twain's counterparts.

How shall we then analyse a statement 'Mark Twain is Samuel Clemens'? Given a few reasonable presuppositions (e.g., that we are talking about the famous American writer using his real name and his pseudonym), the statement is true in the actual world. Because it is a statement of identity between two singular referring terms, it should be true in every possible world where the referent exists. At this point, however, we should take into consideration some further features of counterpart theory. As we discussed in section 3.1.2, the counterpart relation is nontransitive, nonsymmetric, and an individual can have more that one counterpart in another possible world.[10] It need not trouble us now that the relation is nontransitive and nonsymmetric. At the moment, we are interested only in the relation between an actual-world individual and its counterparts, not in the relation between those counterparts and their counterparts, or those counterparts and the actual-world individual. What is troubling is that the actual-world individual, Mark Twain, can have more than one counterpart in another possible world. Which one of them should the name then refer to? A rigid designator should refer to the same individual in every possible world. There is no provision here for an individual splitting in two or more. Within the theory of counterparts we cannot guarantee a unique referent of a singular referring term in a non-actual possible world. This is an unpleasant problem and matters get worse still. A counterpart of an actual-world individual, Mark Twain, even though it is by definition the thing in that possible world that resembles Mark Twain most closely, may still be nothing like Mark Twain in any but a formal sense. Mark Twain's counterpart could be a wax figure resembling him, a gorilla who managed to write Tom Sawyer, or a park bench. In a framework where 'the same individual' may turn,

---

[10]This is discussed in Lewis, 1979b, 113.

in a non-actual possible world, into all sorts of bizarre things, the notion of rigid designation looses its meaning. And this is a natural consequence of the non-essentialist approach adopted by the strongly realist framework.

We can apply some of these observations to an analysis of universal statements, such as - again - 'All cows are mammals'. We can assume that the belief we express using this statement concerns not only actual-world cows, but also cows' counterparts. However, not surprisingly the counterparts may turn out to be very unlike the cows we know, and, indeed, they may not be mammals at all.,

All in all, if we apply to Lewis's work the strict reading, on which nothing exists in two worlds, the notion of necessity collapses for statements involving individuals. On the more lenient reading, where we allow rigid designators to refer across possible worlds, we get such a distorted notion of rigid designation that it seems of little if any use. It may still happen that a pair of universal statements will pick out the same proposition (by being true in the same possible worlds) but because kinds and mass terms are not assigned any essential properties, these cases will occur at random, defying any attempts of systematic description.

The partial collapse of the notion of necessity on the strict reading leads to a proliferation of necessary propositions. All statements involving individuals turn out to be necessary, even if it be only in one world (the actual world of the individual involved). On this reading, one taking the worldboundedness of individuals seriously, the problem of logical omniscience takes two forms, both of them leading to bizarre consequences. We can assume either that proper names are treated as directly referential, or that they are not. If they are, then they help individuate the propositions in which their referent figure. This makes each statement about an individual necessary in the world where that individual exists, and logical omniscience, while formally present, does not lead to a prediction that an agent believes in any more propositions than the one he asserts. If we do not treat names as directly referential, all propositions of the same form regarding individuals are necessary and we cannot separate them (a statement of Pa and of Pb cannot be told apart). On this

reading, logical omniscience becomes extremely pervasive.

On the more lenient reading of the strongly realist proposal some of the problems of the strict reading dissolve but the notion of an individual (a potential referent of a rigid designator) is so strange that the necessity of statements involving individuals does not correspond to anything intuitive.

We have just observed a very strange behaviour of the notion of necessity within a strongly realist framework. As we emphasised in sections 3.1.9 and 3.4, the prediction that individuals are world-bound is integral to the core of a strongly realist approach. The world-bound notion of individual makes seemingly obviates the need for the adoption of any degree of essentialism should be adopted into a realist framework. We saw here that these two features – world-bound individuals and lack of any form of essentialism – were the main culprits responsible for the partial collapse of the notion of necessity and the critical weakening of the notion of proposition. It is fair, however, to repeat that Lewis developed his possible-world framework to deal with other issues (e.g., relative probability of different counterfactuals), and did not intend it to be used to account for the foundational semantics of any kind of terms. His own account of foundational semantics, as we shall see in the following section, takes a completely different route.

4.1.4 Lewis's Folk-psychological Approach to Foundational Semantics

As we mentioned at the beginning of the previous section, the notion of proposition does not play an important role in Lewis's foundational semantics. In fact, he is critical of approaches that take this route. Lewis himself adopts a causal approach to foundational semantics but it is causal with a bit of difference. Lewis's approach is cautiously internalist in making the wide content derivative of the narrow one, it does not presuppose realist views on science, and it is – not surprisingly – rather critical of Kripke's work. And it is presented as grounded in folk psychology.

In order to get an idea of how Lewis's proposal works, we should

get acquainted with the notion of folk psychology. According to
Lewis,[11] we have a good understanding of how people operate men-
tally. We can think of this understanding as a sort of theory – folk
psychology. Folk psychology is shared tacit knowledge, just like the
knowledge of grammar. It is a powerful instrument of prediction –
we can tell which predictions conform to it and which do not. It has
evolved over thousands of years of close observations of one another.
Its predictions do not always turn out to be correct but we can be
confident that folk psychology is largely on the right track. What
can folk psychology do for the project of foundational semantics?
Lewis says that

> Folk psychology concerns the causal relations of mental states,
> perceptual stimuli, and behavioural responses. It says how
> mental states, singly or in combination, are apt for causing
> behaviour; and it says how mental states are apt to change
> under the impact of perceptual stimuli and other mental
> states. Thus it associates with each mental state a typical
> causal role.[12]

Lewis defines a causal role of a mental state as follows: Assume we
managed to elicit all those tacit principles of folk psychology. Let
us say that whenever M is a folk-psychological name for a mental
state, folk psychology will say that the state M occupies a certain
causal role, an M-role. The meaning of state M is then 'the state
that typically occupies the M-role'. A definition of a causal role of
one mental state involves other mental states: causal roles of mental
states are interdefined.[13] The causal roles of mental states involve
responses to perceptual stimuli. Often, the relevant features of those
stimuli will be secondary qualities, e.g., colours. One cannot spec-
ify these qualities in purely physical terms because that would go
beyond what is known to folk psychology. Yet, if we analyse sec-
ondary qualities purely in terms of the responses they are apt to
provoke, the argument will be circular. That is why, Lewis urges,[14]

---

[11]My account of his position is based mainly on Lewis, 1995, 412-431.

[12]Lewis, 1995, 416.

[13]Lewis, 1995, 416.

[14]Lewis, 1995, 416.

we should say that folk psychology includes folk psychophysics. Folk psychophysics tells us that a pair composed of some secondary quality, say colour, and the corresponding sensation occupies a complex causal role that consists partly, but not always, in the former causing the latter. This should do the trick – we have a derivative role associated with the name of the secondary quality, for example colour, and another associated with the name of the sensation. And jointly this pair occupies a complex causal role.

There is another reason to fear circularity in folk-psychological explanations – it concerns the behaviour that mental states tend to cause. We often describe behaviour in mentally loaded terms as 'action'. To say that John passed the salt to his dining companion is to describe his behaviour but it is a description that presupposes a great deal about how his behaviour serves his desires on basis of his beliefs. In describing even just a salt-passing action we might end up invoking John's beliefs about the presence of his dining companion, the words the dining companion used to ask for the salt, the weight of the salt-shaker, etc. A description that would take all of John's beliefs into account would turn into a folk-psychological description of his whole environment.[15] On the other hand, just like in the case of secondary qualities, a purely physical description of the behaviour would go beyond what is known to folk psychology. Fortunately there is another kind of description, one that is available to folk psychology, that we can use. We can say that when John took hold of the salt-shaker, his body moved in such a way that *if* he had been on the surface of Earth with a salt-shaker placed suitably in front of him, under normal gravity, and his dining companion was within his reach, *then* the trajectory of John's arm would deliver that salt-shaker to a suitable distance of his friend. This is not a description people would usually give, but it is recognisable as correct and it does not involve references to John's or his dining companion's mental states.

When we describe a mental state M as the occupant of the M-role, it is a topic-neutral description since it says nothing about

---

[15]Lewis, 1995, 417.

the kind of state that occupies that role. It may be non-physical or physical, and if physical, it may be a state of neural activity or current on a silicon chip. Just what kind of state occupies the *M*-role is an a posteriori matter. This is an important feature of the folk-psychological approach.

Even if we were able to identify a mental state with an underlying physical state, such identification would be contingent because it would be a posteriori. In other words, Lewis argues against Kripke, science is not a vehicle of discovering necessary a posteriori truths.

> Kripke (1980) vigorously intuits that some names for mental states, in particular 'pain', are rigid designators: that is, it's not contingent what their referents are. I myself intuit no such thing, so the non-rigidity imputed by causal-role analyses troubles me not at all.[16]

The motivation for the non-rigidity of names of mental states (e.g., 'pain') follows an epistemic line: Imagine you are in pain. There is some physical state that causes that state. Now consider a counterfactual situation which is exactly like the actual situation except that some other physical state occupies the cause role of pain. How could you distinguish the two situations? You could not. Kripke would predict that in the counterfactual situation you are not in pain. Folk-psychological approach, on the other hand, assumes that usually we know well enough when we are in pain. The term 'pain' refers to a mental state that plays a certain causal role because it tends to cause certain behaviour and influence other mental states. It would not be pain if it did not.

It is quite possible that there is a variation across worlds as to the states that occupy the folk-psychological roles and deserve the folk-psychological names. In fact, these states possibly vary even within the actual world. If we admit that at least some folk-psychological roles are occupied in animals (e.g., that dogs can be in pain) then there is probably a variation between species. And quite possibly here are variations even within one species, for example in how various tasks are divided between the brain's hemispheres.

---

[16]Lewis, 1995, 419.

Lewis's use of folk psychology may seem to demand too much of something that is assumed to be common tacit knowledge but it could be employed quite successfully to avoid a number of problems that the causal approach to mental states has to face. Lewis's explanation uses on the one hand the notion of the normal connection between a physical state and a mental state, on the other hand argues against the rigid designation for names of mental states. In this it is somewhat reminiscent of Stalnaker's information theoretic account of intentionality, which we presented in section 3.3.5, and which seems rather more attractive. It was, however, in the interest of a fair presentation of Lewis's views, that his folk-psychological explanation of content of mental states should be included here. Having done that, we can turn our attention to an evaluation of the problems a strongly realist possible-world framework faces when applied to foundational semantics.

4.1.5   The Realist Concludes

We saw in section 4.1.3 that a non-essentialist possible-world framework containing worldbound individuals does not give rise to an interesting notion of proposition. This is because the notions that underlie the notion of proposition as we discussed it – the notion of necessity, the notion of what it takes to be a particular individual - are not built into the framework in any but the most formal sense. In Lewis's work, the notion of proposition is not intended to be applied to an account of belief. We saw in the previous section that Lewis's own account of foundational semantics takes quite a different direction.

We can now perhaps hypothesise that an interesting notion of proposition relies on a notion of trans-world individual, which is in turn based on some kind of essential properties. By adopting some essential properties into the framework one can perhaps derive a more interesting notion of proposition but the problem of logical omniscience will be still alive and kicking. Just how this happens is what we shall investigate in the following section.

4.2   Propositions in Kripke

We have already touched upon various issues connected with propositions in Kripke's framework in previous chapters. We discussed the source of necessity of various kinds of a posteriori statements in section 3.2.8, and briefly described the distinction between the epistemic and the metaphysical level with respect to statements, truths, and states of affairs (in section 1.7.3). We can apply some insights from those sections to our current problem.

It is well known that in his *Naming and Necessity*, Kripke outlines the difference between epistemic and metaphysical distinctions with respect to statements.[17] We should note that although Kripke uses the term 'statement' fairly consistently, the way he uses it, and the properties he endows it with are basically the same that apply to the notion of proposition. The questions we can and should ask about statements in Kripke are such that we can quite confidently, at least for the time being, treat his notion of statements as if he used the notion of proposition.

The distinction of a priori versus a posteriori is described as epistemic, having to do with the way we came to know the truth of the statements, while the distinction necessary versus contingent has to do with the metaphysical status of the affairs a statement describes. An existing state of affairs is necessary if it could not have been otherwise, that is, if it holds in every possible world, and contingent if there is a possible world where it does not hold. This well-known characterisation does not actually tell us very much. In order to know whether some statement is necessarily true one has to know whether the state of affairs it describes is necessary. In order to know that, one has to know whether the property it ascribes to individuals or kinds involved in the statement obtains necessarily, and that in the end relies on the notion of necessary properties.[18] Whether something is necessarily this or that way depends on the

---

[17]Kripke, 1980, 32-38.

[18]This does not apply to statements of identity between individuals. It seems that we know a priori that if two proper names refer to the same individual, then a statement expressing it is necessary.

way the world is, it is a metaphysical concept and science is supposed to be the tool of unravelling the texture of reality. These are all claims we dealt with when discussing Kripke's work previously. While in themselves these claims may well be controversial, claiming that Kripke indeed subscribes to them is not. We shall now look at the implications these claims have for the issues we are currently investigating: the problem of logical omniscience, the individuation of content, and, more generally, the suitability of the proposition-based approach to content for a foundational-semantics account of the behaviour of proper names.

Firstly, we should establish, just as we did in Lewis's case, whether Kripke ever intended statements or propositions to play an important role in foundational semantics. Just as in Lewis's case, in Kripke's case the answer is not difficult to come by: to Kripke, statements are the carriers of content.[19] It is through statements, by expressing propositions, that people communicate their beliefs and desires. Of course, there are more speech acts than just assertion but Kripke focuses on statements as primary instruments of sharing information.

In Kripke's framework, statements, or propositions, are in a rather good shape: his possible-world framework, unlike Lewis's has a strong concept of necessity as well as a notion of a trans-world individual. We can derive a meaningful notion of a necessary proposition, and the problem of logical omniscience rears its head once again. Kripke distinguishes between the epistemic and the metaphysical aspect of statements, but does that help with respect to the problem of logical omniscience? A speaker may know the truth of a statement in an a priori or a posteriori manner, but the object of his knowledge, the proposition believed, may still be individuated only by its behaviour across possible worlds. And if that is so, then

---

[19] For example in Kripke, 1979, 241, when introducing the Pierre puzzle, Kripke says that on his theory as presented so far, "Whether a sentence expresses a necessary truth or a contingent one depends only on the proposition expressed and not on the words used to express it." and just a little further on the same page he admits that "The situation would seem to be similar with respect to contexts involving knowledge, belief, and epistemic modalities."

there is still only one necessary proposition, one that takes the value 'true' in all possible worlds.

Perhaps, though, we should have a second look at the reasons that lead Kripke to use the term 'statement' rather than 'proposition'. In *Naming and Necessity*, Kripke says that "...it is very far from being true that this idea (that a property can meaningfully held to be essential or accidental to an object independently of its description) is a notion that has no intuitive content..."[20] This turns out to be very important, in fact, one of the main lessons of *Naming and Necessity* is that individuals, and, in general, all referents of rigid designators, have necessary (essential) and contingent (accidental) properties regardless of the description under which they are presented. A statement featuring an individual or a kind – and all statements that Kripke deals with fall under this description – is true or false depending on whether the referent or referents have the property ascribed to them contingently or necessarily or not at all (again, with the exception of statements of identity between proper names). The point is that an individual or a kind can be traced through possible worlds and that it can be found out whether a statement describes it truly or not. The 'statements' of Kripke's work turn out to have more structure than propositions because their truth and falsity depends on the referents of the rigid designators involved in them. While propositions are fully extensional, statements have some structure.

These propositions with added structure, propositions where it matters what individual or kind they are about, are called 'singular propositions'. Kripke does not use that term very much, if at all, but Kaplan does and we talked about them at length previously.

To what degree does the introduction of structure into statements make it easier to account for their use in intentional contexts? We should, after all, keep in mind that while the truth of statements can be known either a priori or a posteriori their truth-value depends, Kripke says repeatedly[21] on metaphysics, which has nothing to do

---

[20]Kripke, 1980, 41.
[21]F, Kripke, 1980, 34-38.

with epistemology. Endowing statements or propositions with this limited kind of structure I have just described alleviates the problem of logical omniscience but does not make it go away. Kripke can capture that believing that Harry is human is different from believing that John is human but cannot distinguish between believing that John is human and that he is an animal or that he has a heart and that he has kidneys. And that is not all. Even if Kripke did not have a problem with necessarily co-extensive properties, there would be a problem concerning ascribing properties to individuals that is grave enough to cause one many a sleepless night. The problem is known as the Pierre (or Kripke's) puzzle.

4.2.1   Kripke is Puzzled About Pierre

We have just pointed out that even after introducing a degree of structure into statements, Kripke still has a problem with co-extensive properties. Clearly, of some properties we do not know whether they are co-extensive or not. It may therefore happen that we use two different sentences to express the same proposition without being aware of it (e.g., if we say that Jimmy has a heart and that Jimmy has kidneys).[22] This consequence of our ignorance of some necessary properties can translate into ignorance of the mutual relation between necessary a posteriori statements. However, according to Kripke, science can and will, given enough time, find out which properties are necessary and which are contingent of various individuals and kinds. This much we have shown (in section 3.2.9). Hypothetically, we could one day find ourselves in a situation where we would know, in all cases, which properties are necessarily co-extensive. A full knowledge of metaphysics would be available. But would this be a situation where we have a full access to a posteriori necessary statements? It seems that this question has to be answered in the negative.

The problem that Kripke – and many others – face has to do with

---

[22]These two sentences are notoriously claimed to be necessarily co-extensive. I have not checked that claim myself and am happy for the moment to rely on the folklore.

the apparent failure of substitutivity of .coreferential proper names in belief contexts. This problem plagues many theories, also theories that are quite different from Kripke's, and we shall briefly mention some of them later, but let us now look at the way Kripke presents it.[23] In this form it is known as the 'Pierre puzzle' or 'Kripke's puzzle'.

Kripke is an advocate of a Millian view of the reference of proper names, according to which

> . . . a proper name is, so to speak *simply* a name. It *simply* refers to its bearer, and has no other linguistic function. In particular, a name does not describe its bearer as possessing any special identifying properties.[24]

We can define this view as follows:

> (**M**)illianism: The meaning of a name is exhausted by its referent.

As Kripke points out, if names are Millian, they should be transparent in the following sense:

> . . . if a strict Millian view is correct, and the linguistic function of a proper name is completely exhausted by the fact that it names its bearer, it would appear that proper names of the same thing are everywhere interchangeable not only *salva veritate* but even *salva significatione*.[25]

A weaker version of this view – one not requiring an identity of meaning, whatever that on the Millian view may be – has been called 'Shakespeareanism'.[26] We can summarise that as follows:

> (**S**)hakespeareanism: Codesignative proper names are inter-substitutable in other expressions *salva veritate*.

---

[23]Kripke, 1979.

[24]Kripke, 1979, 239-240. Kripke commits himself to Millianism in other places as well, for example in Kripke, 1980, 24-26.

[25]Kripke, 1979, 240.

[26]The name was introduced by Peter Geach (Geach, 1968, 165) after quoting Shakespeare's line 'A rose / By any other name / would smell as sweet.'

If names are Shakespearean – and Kripke seems to be committed to that view – they should be substitutable even in contexts involving knowledge, belief, and epistemic modalities. This consequence has been recognised as problematic for a very long time.[27] We can easily imagine a competent speaker who will sincerely assent to 'Cicero denounced Catiline' but not to 'Tully denounced Catiline'. The perceived failure of substitutivity of proper names in belief contexts has been seen as an argument against a Millian theory of names.[28]

Kripke's treatment of the topic stands out because it shows that a puzzle about names in belief contexts can be generated without presupposing a Millian view. It can be derived using much weaker and more general principles that seem rather hard to reject.

We shall assume that a speaker is competent but not omniscient, sincere, reflective, and not conceptually confused.[29] Two general principles are used in deriving Kripke's puzzle. Firstly, there is the disquotational principle, ($\mathbf{D}_{\text{English}}$), stated as follows:

> ($\mathbf{D}_{\text{English}}$): If a normal English speaker, on reflection, sincerely assents to 'P', then he believes that P.[30]

Here 'P' can be replaced by any appropriate English sentence that lacks any indexical or pronominal devices or ambiguities that would ruin the intuitive sense of the principle. Of course, we can fashion analogous principles in and for other languages, including French.

Secondly, we presuppose a principle of translation, ($\mathbf{T}$):

> ($\mathbf{T}$)ranslation: If a sentence of one language expresses a truth in that language, then any translation of it into any other language also expresses a truth (in that language).[31]

---

[27] See for example Frege, 1893, re-print in English 1952, footnote 2.

[28] The example and the criticism has been famously presented in Quine, 1960, 141-146.

[29] A speaker is sincere if she does not consciously say something she believes to be false. She is reflective if she also takes pains to present her beliefs accurately.

[30] Kripke, 1979, 248.

[31] Kripke, 1979, 250.

The Pierre puzzle is well known, so let me just summarise it briefly.[32]:
Let us suppose that Pierre is a normal French speaker who does not
speak a word of English. In the course of his life he has heard about
London (which he, being French, calls 'Londres'). On the basis of
what he heard, he came to the conclusion that London is pretty.
Therefore, he says, in French, 'Londres est jolie', that is, translated
into English, 'London is pretty'. We can therefore conclude on the
basis of ($\mathbf{D}_{\text{French}}$) and ($\mathbf{T}$) that:

> (1) Pierre believes that London is pretty.

Later, Pierre moves to London, to one of its not-so-attractive parts,
and learns English purely by exposure. He learns, among other
things, that the city he lives in is called 'London', and he comes to
the point where he willingly assents to 'London is not pretty'.
    We can then conclude, using ($\mathbf{D}_{\text{English}}$), that

> (2) Pierre believes that London is not pretty.

Pierre, living in London, has no inclination to assent to 'London is
pretty'. He did not learn, however, that 'London' and 'Londres' are
two names of the same city, and is therefore still likely to assent to
'Londres est jolie', that is, that London is pretty. The question is:
does Pierre, or does he not believe that London is pretty?
    If 'London' and 'Londres' have the same semantic value (just
like the notorious 'Hesperus' and 'Phosphorus'), we should conclude,
using ($\mathbf{D}_{\text{French}}$), ($\mathbf{D}_{\text{English}}$) and ($\mathbf{T}$), that Pierre thinks both that
London is pretty and that it is not pretty.
    It seems impossible, Kripke goes on to say, to deny that Pierre
once believed that London is pretty – he was in the same situation
as any monolingual Frenchman who believes that London is pretty.
We also cannot say that now that Pierre lives in London he does not
believe what he once did because he did not change his mind. We
cannot deny Pierre's belief that London is ugly since his command
of English does not set him apart from his neighbours in London.
Thus we seem forced to conclude that Pierre holds both that London
is pretty and that it is not pretty.

---

[32]On basis of Kripke, 1979, 254-255.

> We may suppose that Pierre, in spite of the unfortunate situation in which he now finds himself, is a leading philosopher and logician. He would *never* let contradictory beliefs pass. And surely anyone, leading logician or no, is in principle in a position to notice and correct contradictory beliefs if he has them.[33]

But it is clear that Pierre, as long as he does not know that 'London' and 'Londres' name the same city, is not in a position to revise his contradictory beliefs. He lacks information, not logical acumen.

A similar puzzle can be constructed using just one language, and relying only on the principle of disquotation. In this case, Pierre learns that 'Paderewski' is a name of a famous pianist who lived at the beginning at the 20th century. Knowing that, he will assent to 'Paderewski had a musical talent', and we can infer that

> (3) Pierre believes that Paderewski had a musical talent.

Later he learns of someone called 'Paderewski' who was a Polish nationalist leader. Pierre has a low opinion of musical talents of politicians, and concludes that Paderewski had no musical talent. We can then infer that

> (4) Pierre believes that Paderewski had no musical talent.

From there, it is but a short step to conclude that

> (5) Pierre believes both that Paderewski had a musical talent and that Paderewski had no musical talent.

If this is the case, we seem forced to conclude that Pierre holds contradictory beliefs, and is therefore not rational.

In both the London/Londres and the Paderewski case, we are compelled to believe that a supposedly rational person holds contradictory beliefs, and is therefore not as rational as originally supposed. This conclusion seems unpalatable yet the reasoning is simple and convincing enough. It seems that in order to avoid the puzzle,

---

[33]Kripke, 1979, 257.

we should give up one of the seemingly innocuous premises on which the puzzles are built.

As it is, the Pierre puzzle seems bad enough to give one dizzy spells. The primary cause of its apparent intractability lies in the difficulty associated with finding the culprit step or the unwarranted assumption in the reasoning. That is why responses to the puzzle differ widely in their interpretation of what the puzzle is 'really' about. It can be seen as challenging Kripke's theory of reference of proper names – but we shall show that other theories of reference are vulnerable to it as well. It can make us re-think our views on translation practices or habits of reporting belief. It raises questions about our notions of rationality and content. Approaches to the puzzle differ in their direction as well as depth, and the following attempts at dissolving the puzzle represent, to some degree, the variety of approaches that have been taken.

4.2.2   Naive Contextualism

Contextualism is an approach that attempts to dissolve Kripke's puzzle by challenging the principle of disquotation. It has been championed by, among others, Joseph Moore.[34] The main strategy of contextualism is to show that Kripke cannot exclude from his disquotational principle sentences that contain "indexicals or pronominal devices or ambiguities that would ruin the intuitive sense of the principle"[35] and still make his argument.[36]

To characterise what goes wrong in the Pierre puzzle, a notion of *misdisquotation* is introduced. Misdisquotation is supposed to arise from an oversimplification of the relation between a person's act of linguistic assent and her states of belief. The aim is to show that Kripke 'misdisquotes' Pierre. It is alleged that ambiguities and indexicality can be present in a sentence implicitly, and that the best way to deal with it is by 'contextualising' Kripke's disquotation principle. A contextualised disquotational principle looks as follows:

---

[34]In this section, I base my description of contextualism on Moore, 1999.
[35]Kripke, 1979, 249.
[36]Moore, 1999, 340.

> (**DC**): If (i) a normal, English speaker, x, sincerely and re-
> flectively assents to 'P', (ii)'P' expresses content C in the
> context of assent, and (iii) 'P' also expresses content C in
> the context of attribution, then 'x believes that P' is true in
> the context of attribution.[37]

Principle (DC) is supposed to ensure that we attribute to an agent
a belief in a sentence only if the content of that sentence does not
change between the context of assent and the context of attribution.

Moore characterises content rather loosely as "sets of possible
worlds or states of affairs, perhaps," and then adds: "suppose also
that we take a speaker's assent to a sentence as an expression or
endorsement of the content that sentence has as a sentence in the
shared public language."[38]

It is argued that even a sentence that does not contain any overt
indexicals or ambiguities can change its content between contexts
without the believer changing her mind. Moore provides some exam-
ples of sentences that can change their truth-value depending on the
context. For example, a shift in the implicit domain of quantification
may lead to misdisquotation. Imagine Jenny looking around a party
and declaring 'Everyone is drunk.' She would be misdisquoted if we
said later, in a context of looking for a taxi to take her and her party
safely home, 'Jenny believes that everyone is drunk.'[39] Ambiguity
too can lead to misdisquotation. We can construe an example where
'bank' becomes actively ambiguous in the sentence 'Jones believes
that Smith puts money in banks.'[40] To prevent misdisquotation, the
following principle of contextualism is proposed:

> **Contextualism**: It is possible that a belief sentence con-
> taining no obvious indexicals (or other explicit parameters)
> is true of a believer relative to one context, and – simulta-
> neously, without contradiction, and with no change in the

---

[37]Moore, 1999, 341.
[38]Both quotations are at Moore, 1999, 343.
[39]That is, everyone including potential taxi drivers.
[40]Both examples are at Moore, 1999, 348.

> believer – false of that believer relative to a different con-
> text.[41]

Not surprisingly, Moore goes on to argue that we should not dis-
quote Pierre when he airs his opinions about Paderewski. Saying
'Pierre believes that Paderewski has musical talent' is an example
of misdisquotation.

Moore's contextualism is neutral about how one should seman-
tically explain the ways in which a context affects our use of names
in belief reports.[42]   Allegedly, we are free to invoke neo-Fregean
senses (following for example Forbes[43]), we may refer to names in
a Language of Thought (as Fodor does[44]), or we could employ the
apparatus of shifting 'relevant alternatives' (using Stalnaker's appa-
ratus[45]). A contextualist only has to claim that there is some such
class of semantic entities (like neo-Fregean senses), and that on some
occasions more than one of its members is associated with a unique
actual individual. The nature of those semantic entities is to be de-
termined by a theory of belief reports. Until that is done – and doing
it is not within the scope of Moore's account – we can neutrally refer
to those entities as 'guises'. We can appeal to them in explaining
what the world is like according to Pierre or other puzzled subjects,
and remain Millian. It is not quite clear, Moore admits,[46] whether
the guise-like entities can perform the semantic job that is expected
of them and be, at the same time, epistemically and metaphysically
acceptable. Guises are invoked to help us understand why subjects
do not always know who their *de re* beliefs are about. They are
essential the contextualist explanation:

> In treating Kripke's example, the contextualist will distin-
> guish contexts not only by which individuals are in a domain

---

[41]Moore, 1999, 347.
[42]Moore, 1999, 349.
[43]Forbes, 1990.
[44]Fodor, 1987.
[45]Stalnaker, 1981.
[46]Moore, 1999, 349.

of discourse, but also by which guises are relevant and available for the task of reporting Pierre's beliefs.[47]

In a context where two guises have the same salience, a contextualist can withhold an answer or else say that a sentence containing them does not have a truth-value. According to this approach, it is true that Pierre believes that London is pretty, and also true, albeit in a different context, that Pierre believes that London is not pretty but there is no context in which both sentences are true. Pierre's rationality thus seems redeemed.

This account of the Pierre puzzle seems vulnerable to several kinds of criticism. Firstly, it is grounded in a rather exacting notion of rationality, according to which

> For any rational, non-compartmentalised subject, x, and any sentence 'S' it is not true to say in a fixed context 'x believes that S and x believes that $not - S$.'[48]

Moore does not explain what 'being non-compartmentalised' means but we could infer that it presupposes that a rational agent is capable of accessing all her beliefs and making sure none contradict each other. The opposite, a situation where an agent holds contradictory beliefs, is described as a form of irrationality on the part of the agent, and it is not, according to Moore, to be taken lightly. He says: "...the irrationality might be as mild as insufficient attention or carelessness in cognitive housekeeping, or it might be some more significant breakdown in mental functioning."[49] This conception of rationality, while endorsed by Kripke,[50] has been under attack,[51] and would need a good defence if it is to be sustained. Moore, however, seems quite comfortable about taking it for granted.

Secondly, the principles quoted above – the contextualised principle of disquotation (**DC**) and the principle of contextualism – are not supported by a theory of content and a theory of meaning. The

---

[47]Moore, 1999, 349.
[48]Moore, 1999, 353, notation adjusted for consistency.
[49]Moore, 1999, 353.
[50]Kripke, 1979.
[51]By, for example Salmon, 1986 and Sosa, 1996.

principle (**DC**) tells us that we must not disquote if the content of the assertion changed, but then we should have the means of telling when this happens, and the contextualist approach, at least in the form presented by Moore, does not provide these. Saying that content can be characterised as sets of possible worlds or states of affairs does not count as providing a theory of content. Thus while the idea behind the (**DC**) principle seems plausible, the lack of elaboration on the notions used in it is crippling.

Thirdly, the examples Moore uses to motivate his claim that the content of a sentence can change even if the sentence does not contain *overt* ambiguities or indexicals are not convincing. One of the examples I quoted here contains the word 'bank' – a notorious and overt example of ambiguity, another example contains a quantifier that changes its domain.

Fourthly, it is certainly tenable to admit that sometimes we refer, as it were, 'under a guise', but unless it is specified whether a guise becomes a part of the content, we have not learnt very much. Moore's main idea seems to be that we should not disquote a sentence if its content changes. We could then infer that if this approach is to help us with Kripke's puzzle, a proper name (like 'Paderewski' or 'London') can be responsible for a change in content. Therefore, whatever version of 'guise' we adopt, it has to be semantically relevant. If this is the case, it is hard to see how Moore can claim that he remains Millian.

Finally, the basic strategy of contextualism (at least as it is presented by Moore) – the banning of disquotation if the meaning of a sentence changes between relevant contexts – could be interesting if its repercussions were further explored. The next approach is an example of this.

4.2.3   The Hidden Premise

David Sosa's treatment of the puzzle[52] has been influential, and deserves attention. As we saw earlier, Kripke's aim was to show that

---

[52]Sosa, 1996.

the traditional argument, concerning Cicero and Tully, should not be seen as arguing against Millianism or against thesis (**S**), because an analogous argument can be made without using substitution. Sosa's aim is to show that even the analogue involving Paderewski invokes Millianism in a covert way. Sosa's aim is not to refute Millianism (though he offers an alternative), it is rather to show that Kripke's argument should not be seen as supporting Millianism because it presupposes it.

Sosa claims that when examining the Pierre puzzle about London,[53] one may be tempted to reject the principle (**D**). But disquotation is something we commonly rely on in practice. We may find it questionable or suspect but whatever its merits, something like a principle of disquotation describes our everyday practice of reporting beliefs. Therefore, if we rejected (**D**), we would have to find some other principle to replace it. From a pragmatic viewpoint then, trying to undercut the principle of disquotation seems inadvisable.

The appearances also go against blaming the principle (**D**): the traditional Fregean puzzle, involving Cicero and Tully, seems to point to the substitution (and, consequently, some form of Shakespeareanism) as the source of the trouble, and Kripke's puzzles look analogous to the Fregean puzzle. This similarity is something that Kripke himself is keen to note when describing the aims of his 'Puzzle About Belief', saying

> We will also give an example,..., to show that a form of (Frege's) paradox may result within English alone, so that the only principle invoked is that of disquotation (or, perhaps, disquotation plus *homophonic* translation). It will intuitively be fairly clear, in these cases, that the situation of the subject is 'essentially the same' as that of Jones with respect to 'Cicero' and 'Tully'.[54]

(The principle of disquotation is indirectly connected with the principle of translation. Disquotation is used in belief reports, and translation is used when we want to make the utterances and beliefs they

---

[53] The following arguments are taken from Sosa, 1996, 385-386.
[54] Kripke, 1979, 253.

express available in a language different from the one in which they
were expressed. In both cases, what gets reported is the content of
an utterance – we translate the meaning of what is said (this is es-
pecially obvious in the translation of idioms), and disquote likewise.
In that sense, the two principles are connected.

Moreover, Kripke-style cases can be re-created without using ei-
ther the principle of disquotation or that of translation. David Sosa
reports[55] that he used to think it an interesting coincidence that
there was a well-known politician named John Glenn, since there was
also a famous astronaut of the same name, who, Sosa had thought,
died on a later mission. Now imagine Sosa's situation before he dis-
covered his error: he would have attributed to *himself* the beliefs
that John Glenn has been in space and that John Glenn has never
been in space. What goes on here is not an inference to beliefs from
assertions. It is a direct recollection, and – just like the Paderewski
case – it does not invoke the principle of translation. The point
Sosa makes is that it seems useless to try and stop Kripke's exam-
ples before they reach the stage when, on the basis of the agent's
assertions, beliefs that x is F and that x is not F are attributed
to the agent.[56] It seems then only natural to admit that neither
disquotation nor translation are to blame for Kripke-style problems,
and focus on what is left.

It may be helpful to think about the kind of situations that do not
allow a creation of a Kripke-style analogue. Imagine Rock, a normal
monolingual English speaker who has never left his small town of
Paris, Texas.[57] Though he himself has never been to France, he
heard of the famous city of Paris. On the basis of what he heard, he
is inclined to say, in English, 'Paris is pretty'. Based on his sincere
utterance, we can thus conclude

(1) Rock believes that Paris is pretty.

His opinion of his hometown is not so favourable. Based on what
he knows he is inclined to assert 'Paris is not pretty'. We can then

---

[55]Sosa, 1996, 384.

[56]Sosa, 1996, 385.

[57]This example is found in Sosa, 1996, 386.

conclude that

> (2) Rock believes that Paris is not pretty.

The non-puzzle is: how should we describe the situation? It might seem that on basis of (1) and (2) we could conclude that:

> (3) Rock believes that Paris is pretty and Rock believes that Paris is not pretty.

> (4) If Rock believes that Paris is pretty and Rock believes that Paris is not pretty, then Rock has contradictory beliefs.

> (5) Rock has contradictory beliefs.

> (6) If Rock has contradictory beliefs then Rock is not rational.

> (7) Rock is not rational.

Of course, the 'Paris' of this case is ambiguous. In claiming (3), we do not attribute to Rock any contradictory beliefs. If we were to describe the logical structure of sentence (3), we would use different constants for each occurrence of 'Paris'. We can only use the same constant to represent a proper name if the name is not ambiguous.

One may argue that this point is obvious. Kripke-style and Frege-style cases cannot be generated with ambiguous names. What Millianism implies is that what is semantically relevant to a name is just its denotation. Therefore, names that are co-designative have the same meaning, and are not, properly speaking, ambiguous. 'London' and 'Londres', just like 'Tully' and 'Cicero' may sound different but their semantic value is the same. The principle that seems to be at work in Kripke's puzzles is this:

> (**H**)ermeneutic: If a name in ordinary language has a single *referent* then it may correctly be represented logically by a single constant.[58]

---

[58]Sosa, 1996, 388.

In effect, the principle (**H**) takes having a single referent to be synonymous with being unambiguous. In arguments involving 'Cicero', 'London', and 'Paderewski', (**H**) seems essential in creating the contradictory results. The principle follows from Millianism (though it does not hold *vice versa*): if the only meaning of a name is its referent, then having a single referent is sufficient to guarantee the propriety of representing a name by a single constant.

Sosa's point is that something like the principle (**H**)[59] is implicitly used in the Paderewski puzzle, which was supposed to be a defence of Millianism in the sense that it did not rely on either substitution or translation. Having circumvented and accepted both the principle of translation and the principle of disquotation, respectively, we are left in a position where the most likely culprit is the principle (**H**). It is, Sosa claims, this principle, (**H**), that makes us characterise Pierre as not rational. Kripke-style cases cannot be reconstructed if the term in question is seen as ambiguous. In making his case in support of Millianism, Kripke had to *presuppose* something like the hermeneutic principle, he had to presuppose that a single referent implies a single meaning. But that was what he set out to defend. This is why Sosa finds Kripke's argument insufficient as a defence of Millianism.

Sosa's reaction is to reject the hermeneutic principle, and adopt a Fregean position, according to which even a term that has only one referent can have numerous senses. If we assume that names have meanings over and above their referents, and if we accept that those meanings can vary even if the referent does not, we can then accept that a name can be ambiguous even when it has a single referent. In this picture, a name that has a number of meanings should be represented by different logical constants, one for each sense. "In short, *Fregeanism can see* ambiguities *to which Millianism is blind*."[60] If a Fregean position were accepted, one could make sense of Pierre believing both that London is pretty and that it is not pretty, and

---

[59] It may be that what is at work is a stronger principle (**H′**): A name in ordinary language may correctly be represented logically by a single constant iff the name has a single referent.(Sosa, 1996, footnote 11, 398).

[60] Sosa, 1996, 389, emphasis in the original.

explain how he can remain a rational agent.

A Fregean could insist that the expressions an agent uses to express his beliefs should not contain any ambiguity. But that would be too strong – we can and do report beliefs that contain ambiguous terms, for example definite descriptions. "However – and this is crucial – *our use of ambiguous expressions in the belief attribution should be disambiguated in a way that* matches *the original use of the ambiguous term by the person disquoted.*"[61] Similarly, we translate ambiguous expressions into our language correctly only if we disambiguate the terms in question correctly.

Sosa's argument is designed to show that in order to produce a Kripke-style case, something like the principle (**H**) has to be presupposed. This creates a problem for Kripke on two counts: Firstly, Kripke intended to show that a Frege-like puzzle can be reconstructed using only assumptions about disquotation and rationality, and now it turns out that he used a Millian premise. Secondly, (**H**) is the weakest part of the argument. Adoption of a Fregean position, that is, a rejection of the principle (**H**), leaves in place the much less controversial premises used by Kripke, and dissolves the puzzle. Sosa's aim is neither to claim that Millianism is indefensible nor to show that Fregeanism is the only way out of the trouble. What Sosa shows is that Kripke's argument does not support Millianism because it presupposes it.

4.2.4 Sophisticated Localism

Localism is a theory developed by Akeel Bilgrami[62] in response to both Frege's and Kripke's puzzle. Its aim is to harmonise two sets of desiderata. Firstly, it seems desirable that beliefs attributed to an agent should make her rational by her own lights, that is, represent a consistent state of affairs in the world as she conceives of it. Secondly, the content of an agent's beliefs should not be separated from her external circumstances in a way that would allow for

---

[61]Sosa, 1996, 394-395, emphasis in the original.
[62]In Bilgrami, 1992.

scepticism about the external world.[63] These two sets of constraints seem to pull in different directions. Depending on which is seen as more pressing, one can see theories as being internalist or externalist in their motivation, or accounting primarily for narrow or wide content.

Bilgrami starts off by trying to account for the internalist intuitions, and then proceeds to explain how the content thus specified (narrow content) relates to the external world. Admittedly, a theory that starts off from the narrow or Fregean notion of content cannot be made compatible with externalism of a Kripkean or Putnamian kind. The task is therefore to find a version of externalism that could fit the bill. The following example[64] illustrates the direction Bilgrami takes.

Imagine two agents. One of them knows chemistry, the other one does not. The difference in their knowledge should make a difference to the way an external substance determines the concept they express with the term 'water'. That is, though it is the same substance, water, that determines their concepts, each agent's concept will be different if their background knowledge differs. Because it is an external substance that determines their concepts, a commitment is made to a kind of externalism about content, but because the way in which an external substance determines the concept is constrained by the background beliefs of the agent, the content thus determined captures the external world as the agent conceives of it.[65]

The main problem with content conceived of along these lines is that it would be too finely grained since it is unlikely that any two agents have exactly the same background beliefs regarding an external substance. Bilgrami addresses this concern by pointing out[66]

---

[63]In my presentation of Bilgrami's views I follow his Bilgrami, 1998.

[64]From Bilgrami, 1998, 597.

[65]One may argue that any reasonable internalist theory should allow for this degree of determination by external circumstances. This would be a valid remark but Bilgrami's point here is to specify the degree to which an externalist element is incorporated into his theory. Whether we call his theory moderately externalist or moderately internalist is just a question of labelling.

[66]Bilgrami, 1998, 598.

that a form of contextuality of content – what he calls 'locality' of content – offers a common-sense solution. The notion of locality is illustrated by the following example. Imagine once again the two agents, one of whom knows chemistry, whereas the other does not. Let us say that the former one is drinking a certain substance from the tap because he is thirsty after a game of tennis. In this context, his background chemical beliefs are irrelevant to his concept of 'water'. In this context, therefore, he shares his concept of 'water' with the latter agent. In fact, in most contexts of thirst-quenching we can abstract from the differences between these agents' concepts of 'water'. In some contexts, however, such as in a chemical laboratory, the differences in background beliefs do come to the foreground.

In the context of the Pierre puzzle, this helps explain how Pierre's concept of 'London' will change as his external circumstances alter and he moves to London. When he lived in France, Pierre shared his concept of 'Londres' with other Frenchmen who have never been there and knew it only from postcards and other media that represented it in a favourable light. When he moved to London, he started acquiring a concept of that place in a similar way in which other Londoners who lived in the same area did. Then he came to the conclusion that London was not pretty. The 'expert knowledge' he does not possess concerns the fact that 'London' and 'Londres' are two names for the same city. In most situations, however, his concept of 'London' will be very similar to that of other Londoners, and the lack of expert knowledge would only come to the foreground if he, for example, met another Frenchman, one who would know that 'London' and 'Londres' are one and the same place.

This approach to belief and content has to address some concerns related to the normativity of meaning and the social aspect of language. It is here that Bilgrami's approach becomes more controversial.

There is a sense in which meanings are not social. One need not attribute to an individual the concepts of his fellows if that individual does not possess some particular expert knowledge that they do possess. An individual can be at variance with the experts, yet even his individual meaning can be seen as public because others can and

do understand idiosyncratic meanings. Thus, Bilgrami argues, the private *versus* public distinction does not need to coincide with the individual *versus* social one. It is possible for a speaker to display publicly her individual meaning (people sometimes use words in non-standard ways and when prompted, explain) just as it is possible to express privately a social meaning. This idea, that an individual's meanings or concepts are not necessarily constituted by the socially constituted reference, prompts the objection that Bilgrami does not account for the normativity of meaning. To illustrate it, we can use an example provided by a well-known advocate of semantic externalism, Tyler Burge.[67] In the picture Bigrami presents, if Burge's arthritis sufferer – let us call him Bert – claims to have arthritis in his thigh, he can be counted as saying something true given what *he* means by 'arthritis'. He can vary from the majority by what he means by that word, and yet is not be counted as being wrong in his use or the underlying concept. This is a consequence Bilgrami accepts.

The application of Bilgrami's localism to belief reports has also attracted some criticism.[68] Let us look again at Bert, the arthritis sufferer. Bilgrami concedes that "we would, in everyday reporting of belief, say 'Bert believes he has arthritis in his thigh'"[69] but he also claims that this common usage *misleads* us about the actual belief content. Ebbs criticises Bilgrami's rejection of our ordinary practice of attributing beliefs. Yet Ebbs himself proposes a distinction between 'concepts' and 'conceptions' of external entities, such as arthritis, where the *concept* of arthritis is logically independent of any individual's beliefs about arthritis and it is only the *conception* that captures them. It may be argued, however, *pace* Ebbs, that what is relevant to understanding Bert is his conception of arthritis, and the same holds for every speaker. Concepts are then introduced only to account for the normativity of meaning and there is actually very little else we can say about them.

Bilgrami's theory may have been inspired by Kripke's puzzle but

---

[67]The original thought-experiment was presented in Burge (1979).

[68]The following point was made by Ebbs, 1998.

[69]Bilgrami, 1992, 74.

his proposal amounts to a new and refreshing view on meaning in general, not just meaning of proper names and natural kind terms. The consequences of his view are many, and I shall not recount them here – it would take us too far from the direction we set out to follow in this chapter.

4.2.5   Keeping Files

There is an approach that has a localist flavour but focuses on proper names – these are the so-called 'dossier' or 'file' theories of proper names.[70] It has been around for some time and exists in numerous forms. The form I shall present here was developed by Graeme Forbes.[71] Like David Sosa, Forbes thinks that in the end the Paderewski puzzle relies on a Millian premise, and that if we want to solve it, we have to give an alternative explanation of the functioning of proper names. The proposal he advocates is a modified Fregean position. According to Forbes,

> ...when we receive what we take to be *de re* information which we have an interest in retaining, our operating system may create a locus, or dossier, where such information is held; and any further information which we take to be about the same subject can be filed along with the information we already possess. More precisely, the system files what I call "classified conditions"; a condition stands for something that an object can satisfy, and the classifier is what specifies the subject's attitude to a certain related proposition.[72]

An example of a classifier is 'believed to be true' or 'hoped to be true'. The role of a name is to identify a file or label a dossier. The

---

[70]This approach has been championed by for example John Perry (1980), and Segal and Larson (Segal and Larson, 1995 and Segal, 2001). The term was probably first used by Garreth Evans (in Evans, 1982, 399). Its analogue in the formal semantics, however, appeared first – I mean the discourse representation theories of as presented for example in Heim, 1988, Kamp, 1981, and Kamp and Reyle, 1993.

[71]In the following presentation of his views I rely on Forbes (1990) and Forbes (1996).

[72]Forbes, 1990, 538.

cognitive significance of a sense of a name then is 'the subject of *this* dossier'. As we learn more about the subject of the dossier, we may delete or add new information. Applying this to the Pierre puzzle, we may say that there is no confusion as to what Pierre's beliefs are. He has two files under two homonymous names 'Paderewski'. The problem is in figuring out how to properly report his beliefs. Forbes approaches this problem by saying that this may be a case of *expressive inadequacy*.[73] The way we tend to deal with it is by enriching our language by *qualified* names. We shall thus speak of Pierre's beliefs regarding Paderewski the pianist and those regarding Paderewski the politician. In this way, we shall, in our report, do justice to Pierre, save him from irrationality he is patently not guilty of, and get out of the puzzle.

One may object that both Sosa and Forbes owe us a better and more detailed defence of the positions they take in order to solve Kripke's puzzle. One possibly troubling objection is that different agents attach different senses to the terms they use, which leaves us in a difficult spot if we want to explain how is it possible that those terms are used in a successful communication.

4.2.6   What Pierre Taught Us

Beside the obvious things, like how to say 'London' in French, we learned quite a few lessons from Pierre. My selection of responses to Pierre was not intended to be fully representative because some of the most influential accounts of the puzzle take the form of full-fledged theories of proper names and belief[74] the pursuit of which would lead us away from the direction we chose.

Still, there are some observations we can make on the basis of the literature we reviewed. First of all calling it a 'puzzle' is somewhat misleading in that it suggests that there is a single good solution that would become apparent once we got all the pieces in their right place. We have seen, however, that Kripke's scenarios unearth a whole cluster of problems that touch upon a number of basic problems of

---

[73]Forbes, 1990, 561.
[74]I have in mind Salmon, 1986, and Dummett, 1973b.

philosophy of language and mind – what is a proper name, what are the objects of our beliefs, what is the connection between belief and a public expression of it, what are the criteria of rationality, and what are the criteria of understanding a speaker correctly? The simplicity and the intuitive appeal of Kripke's examples give all these questions a renewed urgency.

We have seen that approaches that attempt to prevent the puzzles from developing are not very attractive in the end. To claim that the principle of disquotation should not be used suggests a disregard for a common practice. It is unappealing because what if not our actual linguistic practices should be our guide for analysing natural language.

We should bear in mind that convincing Kripke-style examples can be constructed without any recourse to the principle of disquotation (witness the John Glenn case) or the principle of translation (as in the Paderewski case). In a case like that involving John Glenn, we cannot place blame on the reporting practice because it is the speaker himself who analyses his beliefs and utterances, and we cannot find fault with a translation process because no such process is going on. An analysis of this kind of examples leads us thus to an investigation of the content of speaker's utterances and his beliefs. That does not mean, though, that there is nothing interesting to be said about the varieties of the quoting practice. Investigations in this direction have led to some interesting results, though I have not mentioned them here.[75]

It would seem that accounts that focus on the principle of substitution and those that focus on content can be, more or less, phrased in terms of each other. A satisfactory characterisation of the content of speaker's utterances would give us the means of ensuring that we disquote and substitute not only *salva veritate* but also *salva significatione*. The 'dossier' theories of proper names and localism take this direction.

The difference between the 'dossier' theories and localism illus-

---

[75]I have in mind the polemic, which started mainly in response to Davidson's 'On Saying That', (Davidson, 1968). Lepore and Capellen's article on quotation (Lepore and Cappelen, 1997) is definitely worth reading in this context.

trates another interesting point: some authors who take a theory of content to be the main challenge posed by the Pierre puzzle focus on a general theory of content (e.g., localism), others deal primarily with a theory of proper names (e.g., the 'dossier' theories). Both approaches can have interesting results but it should be born in mind that puzzles similar to those presented by Kripke can be generated using demonstratives (I have in mind Quine's Ortcutt case). This is not surprising because both proper names and demonstratives have been claimed to be directly referential. What the approaches dealing with content have in common is their preoccupation with the issue of externalism: they try to see to what extent the external circumstances we find ourselves in determine the content of our thoughts and utterances.

In the responses to the Pierre puzzle I presented in previous sections, I have not dealt explicitly with the issue of standards of rationality. It should be mentioned at least here. Kripke and a number of other authors[76] claim that a speaker who is caught holding contradictory beliefs should be seen as irrational and somehow in deep trouble. Yet it would seem that having contradictory beliefs is a frequent predicament. Just how frequent depends on what theory of content we champion. If, for example, we hold that the content of utterances is best characterised by fully extensional, non-structured propositions, being wrong about a single necessary proposition (even by thinking, for example, that the statement that expresses it is contingent) would be enough to condemn a speaker to irrationality. If, in using singular propositions, we introduce a degree of structure to characterise content, the situation improves marginally. As long as we hold that the content of an utterance is characterised by sets of possible worlds where the situation described by the statement holds, we cannot make a difference, on the level of content, between various ways of expressing necessary propositions. In the case of singular propositions, this consequence is limited by the fact that a proposition includes the individual figuring in the statement. Yet even necessarily equivalent statements regarding just one individual

---

[76]For example Sosa, 1996, and Moore, 1999 just to name a few.

can be expressed in more ways that the speaker is aware of. It would seem that we have neither the logical acumen nor the knowledge required to ensure that we do not hold mutually contradictory beliefs. I suspect that being inconsistent is commonplace even among logicians. If that is so, how come it does not get us in trouble all the time? Some philosophers, e.g., Dummett,[77] have argued that we do not get in trouble because we do not actively hold all of our beliefs all the time. We call on our beliefs in the relevant contexts. Our beliefs and our knowledge are to some degree compartmentalised, and we feel the need to resolve an inconsistency only when we become aware of it.

Kripke claims that ". . . surely anyone, leading logician or no, is in principle in a position to notice and correct contradictory beliefs if he has them."[78] Yes, we correct contradictory beliefs when we notice them, but are we in principle in a position to spot all of our contradictory beliefs? To do that, we would have to be able to run through all of our beliefs and check them. Unlike Kripke, I do not think that is possible.

Finally, in the background of both Kripke's article and the responses it provoked there is a network of assumptions about the individuation of proper names. It is a difficult topic that deserves special attention.

4.2.7   One Name or Two

A Fregean strategy à la Sosa and some internalist approaches to the puzzle imply that we should individuate proper names in a different way than Kripke does. This could be criticised because it amounts to a change in the basic assumptions of the puzzle. On the other hand, spelling out exactly how Kripke individuates proper names is not as simple as it may seem. What matters here is that if it turned out that Kripke's assumptions in some way presuppose Millianism (which he sets out to defend) then they should be up for discussion,

---

[77]Dummett, 1973b.
[78]Kripke, 1979, 257.

and their modification or rejection should be seen as a legitimate
move in a criticism of the puzzle.

In trying to establish what picture of individuation of proper
names Kripke presupposes, we have several clues to follow. Firstly,
Kripke defends a Millian thesis according to which the semantic
relevance (or the meaning) of a proper name is its denotation, and
not its connotation. Closely connected with this thesis is the claim
that names are rigid designators, and the descriptions we may use
to fix the name's reference do not become a part of its meaning.

> What may be the case is that we *fix the reference* of the
> term 'Cicero' by use of some descriptive phrase, such as 'the
> author of these works'. But once we have this reference fixed,
> we use the name 'Cicero' *rigidly* to designate the man who
> in fact we have identified by his authorship of these works.[79]

According to Kripke, the two ways then in which Pierre seems to use
the name 'Paderewski' – 'Paderewski the politician' and 'Paderewski
the musician' – do not amount to a genuine bifurcation of the name.
As long as 'Paderewski' refers to one and the same person, it is just
one name. Within Millianism, we are not in a position to posit an
ambiguity if what we mean by ambiguity is one word having more
than one meaning. This seems to be a consequence of the Millian
view that was intended and argued for by Kripke.

Can we generalise this, and say that when words refer in the
same language to the same object they have the same meaning? If
this were the case, would it imply that 'Tully' and 'Cicero' are the
same proper name because they have the same meaning? Kripke's
position on this point seems to be that while co-referential names are
synonymous they are not identical. On a metaphysical level, though,
the identity between referents of coreferential. names is necessary.
He says that

> . . . whenever '*a*' and '*b*' are proper names, if *a* is *b*,. . . , it is
> necessary that *a* is *b*. Identity statements between proper

---

[79]Kripke, 1977a, 92.

> names have to be necessary if they are going to be true at
> all.[80]

However strong the synonymy between '*a*' and '*b*', we cannot conclude that the two are identical. We can suppose that the reason behind this is that '*a*' and '*b*' are distinct words. So while the function of proper names is, according to Kripke, simply to refer,[81] they are also – at least to some extent – a part of language. Let us see what we mean by this, and what the repercussions are.

In the presentation of the Pierre puzzle, the French name 'Londres' is translated into English as 'London' using the principle of translation for English. The possibility of translating a proper name seems to be an essential part in the derivation of the puzzle. We would not translate into another natural language a term or a formula that is not a part of that language but belongs to some other symbolic system, for example musical notation.[82] It may then seem that proper names are a part of natural language in a similar way in which for example natural kind terms are. If proper names are, indeed, a part of language, it would seem that knowledge of those names is relevant to a speaker's competence in a given language. Of course, no speaker can be expected to be a competent user of all expressions of any language. We certainly would not want to say that an Englishman who does not know what 'catalpa', 'Pripet' or 'Glengarry'[83] mean, in the sense of knowing even just whether these are names of people, artefacts, natural kinds or what not, is less than fully competent in his language.

On the other hand, could someone be a competent speaker and not know any proper names or natural kind terms at all? This question leads us to the issue of a connection between linguistic

---

[80]Kripke, 1977a, 73.

[81]See Kripke, 1977a, 72: "It would, therefore, seem that the function of names is *simply* to refer, and not to describe the objects so named by such properties as 'being the inventor of bifocals'..."

[82]There are different standards of musical notation between which we can translate, but that is not relevant to my point here.

[83]All these expressions are included in *The American Heritage Dictionary of the English language*, Morris, 1981.

competence and world knowledge. It is hard to imagine a competent language user whose world knowledge is so restricted as to miss all proper names and natural kind terms.[84] People whose language misses too many of those proper names and natural kinds terms one could reasonably expect them to use(those directly relevant to their everyday lives), are either individuals whose language function is compromised by a disease or an injury (e.g., aphasiacs) or people whose competence in other areas of life is limited by some disorder in their broader mental functioning.[85] It is reasonable to expect a competent speaker to be able to use language in such a way that she can influence other speaker's beliefs in the way she intends, and acquire new beliefs that are relevant to her life, that is, that she can make herself understood and understand others. That is why it is reasonable to expect a person to know names of people, animals and substances she is likely to encounter in her environment,[86] and names of places and people around her. There are, however, some substances that are universally present in people's lives, e.g., water. We can then expect every speaker to be able to use their names.

Language competence and competence in other areas of life are linked. We cannot spell out exactly how much of either is required for someone to be considered a competent speaker but cases of extreme deficiency highlight the connection.

Acquisition of a second language deserves a special consideration. There, at least in the Pierre case, we have a speaker who is already competent in one language. Without making any controversial assumptions about a connection between language and beliefs, we can assume that a speaker is likely to express his beliefs in that language in which he has the best chance of being understood by the persons he communicates with at a given occasion. We might say that for a

---

[84]I believe a similar point could be made about other kinds of expressions but want to keep focus on directly referential terms.

[85]The language abilities of individuals in early stages of language acquisition is, of course, quite a different question.

[86]I suspect that for example the reason why fewer and fewer people know the names of even rather common plants is that their relevance to our lives has diminished.

truly bilingual speaker the competence in both languages should be the same, meaning that the speaker should be able to express the beliefs he has as a speaker of one language just as well in the other one. If a native German speaker has a belief which he expresses by saying 'Schnee ist weiss', and his English is as good as his German, then he would also assent to 'Snow is white'. If he believes that Venedig ist schön, he also believes that Venice is pretty.

However, we should not take this line of thinking too far. Our linguistic competence regarding particular terms is limited by our knowledge of the items they name. One may well believe that the black widow is a very dangerous spider yet fail to recognise it in one's bathroom.[87] Pierre's problems with London are so puzzling because, on the one hand, his situation is like the situation with a black widow (where we have a competent speaker lacking some particular information), and on the other hand, he fails to map his 'French' beliefs unto his 'English' beliefs even though they are highly relevant to his life in London. His ignorance regarding the relation between 'London' and 'Londres' is so astonishing as to make us doubt his linguistic competence.[88]

There is a difference between proper names and natural kinds terms that should be mentioned. Proper names, unlike natural kind terms, are frequently imported into languages without much further

---

[87]I know a person to whom this happened – his 'pet' spider in the bathroom turned out to be a black widow. Fortunately, it was pointed out to him in time, and he lived to tell the tale.

[88]One could also point out that the very set-up of the puzzle is rather far-fetched. It seems quite unlikely that Pierre got to London from France without knowing where he was going, and that he lived there without coming in contact with the places he knew from postcards or other sources that contributed to his opinion that 'Londres est jolie'. One could, however, modify the puzzle to make it more plausible. Imagine someone, say Jane, who is a keen amateur reader of Roman history. So she forms a very unfavourable opinion of a place in England called 'Durolipons'. Not only was it just a couple of huts huddling together in mud and rain, giving occasional shelter to highwaymen, but the very aspect of the place seemed unfriendly and not fit for habitation. She would never want to live there. Yet then she applies for a job in Cambridge, quite certain that it is a place where she would like to live. We, of course, rightly suspect that Durolipons and Cambridge are one and the same place.

ado. Place names especially are often simply 'hijacked' from their language of origin. However, names of places that have been historically significant for the speakers of some language sometimes receive a name in that language that may be quite different from their original name ('London' and Londres', 'Köln' and 'Cologne', 'Warszawa' and 'Warsaw' for example), and it is good form to use, for example, an English name when speaking English. But what of the names that do not change between languages, for example 'Amsterdam' in Dutch and in English? Is there one name in two languages or are there two homographic names? And what about Amsterdam, New York? Does it have a Dutch name or an English one? I think these are genuinely grey areas where traditions of use dictate the most effective approach, that is, the approach that makes communication run as well as it can.

4.2.8   Conclusion

We saw in section 4.2 that being able to separate the metaphysical from the epistemological status of statements is very important to Kripke's project. This distinction makes it at least *prima facie* plausible that we could know or believe necessary propositions because if those propositions make factual claims about the world they are likely to be a posteriori. The problem of equivalence of necessary propositions persists but is mitigated by adopting the notion of a singular proposition. Singular propositions involve in an essential way the individual or natural kind they are about, and this limits the number of sentences that can be seen as expressing the same proposition.

The Pierre puzzle presents a serious problem for Kripke because it can be interpreted as challenging the presuppositions of direct reference and of rigid designation. To blunt the challenge, an advocate of Kripke's position would have to show that the puzzle can be dissolved without altering Kripke's position on proper names (or, more generally, without altering the account of content that Kripke seems to presuppose). We did not find such a defence of Kripke's position in the literature we reviewed, but that does not mean that

it could not be done. As I noted earlier, the responses to the puzzle presented here are only a sample of the relevant literature that is out there. We could see, however, that there are a number of difficult problems that have to be faced in solving the puzzle. It is important to keep in mind that any solution to the Pierre puzzle that is incompatible with a Millian view of proper names would challenge some substantial claims that underlie Kripke's work. If we conceded that proper names have a meaning that goes over and above their reference and that this meaning is semantically relevant, i.e., that it influences the truth-value of the sentence in which the name figures, we would have to substantially modify the claim that names are directly referential, and that would complicate the applicability of the notion of rigid designation. Kripke does not claim that names do not have any meaning in terms of information value or descriptions attached to them, but he does claim that whatever sense they have is not a part of their semantics. The Pierre puzzle challenges these central claims of Kripke's work.

4.3 Stalnaker and Necessary Propositions

Among the philosophers whose work we have been examining, Stalnaker worries the most about the problem of logical omniscience. He mentions the vulnerability to it as a drawback of his approach in Stalnaker, 1984, and devotes two articles ('The Problem of Logical Omniscience' part I and II[89]) to it. We have outlined the problem at the beginning of this chapter, in section 4.1 but let us now return to it, and see why Stalnaker finds it so troubling.

In section 3.3, we investigated Stalnaker's approach to modality. We saw that according to him, propositions do not reflect the structure of sentences that are used to express them. In fact, on Stalnaker's view, propositions are conceptually independent of language. For a given state of affairs and a set of relevant possible

---

[89]Stalnaker, 1999c and Stalnaker, 1999d.

worlds, propositions tell us where a state of affairs holds and where it does not. They are purely extensional.[90] Their role is to deliver informational content, which is identified with their truth-conditions. This coarse-grained approach to content gives a special urgency to the following questions: When believing a proposition, which of its consequences and applications are believed as well, so to say, by the same token? How is it possible that someone who believes that P can fail to believe that Q even if P and Q are equivalent?

Here is what we shall do in the following sections: In Stalnaker, 1999a, Stalnaker describes what he considers the received view of how beliefs are stored. We shall try to see how it works, and contrast it with the view that Stalnaker himself advocates. Finally, we shall connect the position Stalnaker arrives at (in Stalnaker, 1999a and Stalnaker, 1999d) with other themes from his work, focusing on the Gricean accents of some of his early articles. Our aim throughout the following section will be to examine Stalnaker's approach to belief, and to focus on the way he tries to deal with problems that arise along the way. One of those problems is the notorious problem of logical omniscience. We hope to place Stalnaker's views on belief in the context of his own work and competing theories. Let us start by presenting what Stalnaker considers the 'received view' of storing beliefs.

4.3.1 Storing Sentences in Boxes

The sentence storage model of belief is implicitly assumed by so many philosophers that it deserves to be called 'the received view'.[91] It has been championed, among others, by Gilbert Harman, Jerry Fodor, and Christopher Cherniak.[92] Stalnaker analyses this approach to belief, using its shortcomings to motivate his own proposal. In his analysis, Stalnaker makes an important assumption

---

[90]We shall delve somewhat deeper into Stalnaker's conception of propositions little later, in section 4.3.3.

[91]The presentation in this section follows Stalnaker, 1999c but elaborates on some points, and goes beyond it in others.

[92]Harman, 1973, Fodor, 1981, Cherniak, 1986.

about what he calls the 'ordinary' notion of belief. According to the 'ordinary' notion, belief is supposed to be, at least to some degree, consciously accessible to the agent who has it. For the time being, let us leave open the question whether this assumption is justifiable.

According to the storage model, one's beliefs are determined by a set of sentences that are stored (in either some sort of a mental language or in one's natural language) in one's belief box. To believe that P is to have a sentence saying that P stored in one's belief box. This is an approximation because no one actually claims that *everything* an agent believes is stored in his belief box. The sentence storage model distinguishes between explicit and implicit beliefs. Explicit beliefs are those that are stored in the belief box, while other beliefs, for example some obvious consequences of the explicit beliefs, are said to be believed only implicitly.

There are two kinds of relations that are important to the contents of the belief box. Firstly, there is the relation between the belief box and the external world (relation between beliefs and sentences), and secondly there are mutual relations between the items in the belief box (relation between sentences and entailments between beliefs). The relation between a belief box and the external world is complex because it includes not only an agent's interactions with his physical surroundings but also his relations to other agents, their beliefs, and his conception thereof. Right now, let us focus on the mutual relations between the items in a belief box.

It is important to keep in mind the difference between the contents of a belief box (i.e., sentences of one or another language) and the content of the sentences in it, which is often identified with the propositions those sentences express.

In order to make some sense of the usefulness of the belief box, we need to characterise the content of the sentences in it, their semantic value. This is because, while the sentences are the explicit beliefs, it is what the sentences say, their content, that figures in belief attributions. In a sentence of the form 'x believes that P', the P is *used* rather than mentioned. A semantics for the language in which the sentences are stored, should tell us what 'x believes that P' means, not in terms of the syntax of P, but in terms of the semantic

value of P. Thus even though to believe something explicitly is to store the appropriate sentence, just which sentence this is, is not necessarily clear from a particular belief attribution. When we say that x explicitly believes that P, all we can conclude is that x stores *some* sentence that says that P.

Also, just which entailments characterise the contents of a belief box depends on how we characterise the contents of the sentences in it. In doing that we have to balance two sets of considerations. Generally speaking, if we individuate the contents of sentences finely, there will be less equivalencies between sentences, which means that if we say that x believes that P, there will be less sentences that could stand for P (and be stored) than if we chose a more coarse-grained description. If the contents are individuated too finely, it will not be possible to bring out the similarities between different sentences stored in the belief box – sentences that play the same role in the cognitive economy of the agent who stores them.

If, on the other hand, we individuate the contents coarsely, e.g., by the sentences' truth-conditions, we end up with a model that seems to predict something close to logical omniscience. In such a model, anyone who believes a necessary truth (e.g., a trivial tautology) believes by the same token all necessary truths. On this approach to the individuation of content, if I believe that all bachelors are unmarried, I believe something that is true in all possible worlds within a given domain, and that cannot be distinguished from anything else that is also true in all possible worlds in that domain. And if I believe a contradiction, then I should believe all its consequences, that is, anything. Advocates of this version of the storage model would say, however, that nothing they claimed so far implies that believers have to be able to *recognise* the logical consequences of their beliefs. We should realise that taking this step is tantamount to giving up on explaining the 'ordinary' conception of belief where we take it that agents are capable of recognising at least some consequences of their beliefs. Thus in responding to an obvious problem with a coarse-grained conception of belief, its advocates gave up on trying to explain how belief functions in real agents.

One may wonder what sorts of conditions should be met by the

contents of the belief box. At this stage, when we are interested in what goes on in the belief box rather than in determining what makes particular beliefs correct, it may seem that no conditions need be imposed on what goes into the belief box. This may well be true, but the question remains whether any principles should apply to the relations between the sentences in a belief box.

Some proponents of the sentence storage model argue that sentences in the belief box should meet at least some minimal standards of coherence. The suggestion is that if the contents of a belief box were a complete chaos, it would not count as a belief box and its owner would not count as an agent. Upon reflection we can see that an organisation of the sentences in the belief box would facilitate the creation of complex beliefs – they can be accessed without having to go through too much of the content of the belief box – as well as the creation of implicit beliefs. Perhaps then it would be plausible to adopt some restrictions on the contents of a belief box. However, what the restrictions should be is far from clear. Considerations of minimal rationality may require that one should believe some obvious consequences of the sentences stored in one's belief box but it cannot be required that these consequences be stored. This is because obvious consequences of sentences that *are* in the belief box are just the kind of thing one should *not* have to store since they can always easily be inferred. Harman's principle of clutter avoidance tells us not to overcrowd one's own belief box with trivial consequences,[93] and this is one of the cases where it seems to apply. Even if we considered an idealised agent whose storage capacities and deductive powers are limitless, we would have no reason to assume that her belief box is closed under deductive consequence because such an agent would find it very easy to always infer obvious consequences of her explicit beliefs.

We saw now that in the case of explicit belief, deductive omniscience is not plausible as a constraint on sentences stored in a belief box neither for real nor for idealised agents. Real agents have a limited storage capacity and should therefore avoid storing obvi-

---

[93]Harman, 1986, 12.

ous consequences of their beliefs, and idealised agents do not have
to store clear consequences of their beliefs since they can always eas-
ily infer them anew. It seems therefore, Stalnaker concludes,[94] that
a logic of explicit belief based on the sentence storage model would
reveal nothing about the inferential powers of either real or idealised
agents since it would deal only with the base from which the agents
reason and not with the ways in which they do it. And if this is the
case, the model misses out on an essential aspect of belief.

Let us turn our attention to a more promising topic now: the
storage model's conception of the implicit belief. Within the storage
model of belief, we find at least two basic conceptions of implicit
belief – a broad one, and a narrow one.

On the broad conception, implicit beliefs include all implicit con-
sequences of the explicit beliefs that are stored in the belief box.
This includes all deductive consequences of explicit belief. On the
broad conception, then, assuming classical logic, belief is deductively
closed. However, the proponents of the storage model would say that
we have no licence to assume anything about the speaker's ability to
access the implicit beliefs. They do not claim that implicit belief is
anything like belief in any ordinary sense. All they claim is that the
notion described here is of some interest. On the broad notion, im-
plicit belief is described by an ordinary modal logic. It is important
to note that because implicit beliefs are not assumed to be accessible
to the agent who supposedly has them, this notion of belief tells us
no more about the inferential powers of a believer than the notion
of explicit belief did.

On the narrow conception of implicit belief, something is believed
only if it is 'easily inferable from one's explicit beliefs'.[95] Easy infer-
ences are supposed to include both some easy inductive and some
obvious deductive consequences of the explicit beliefs. In order to
clarify the narrow notion of implicit belief, one would then have to
specify what counts as an easy inference, and that is far from easy.
But let us suppose for a moment that we do have some definition of

---

[94]Stalnaker, 1999c, 249.
[95]Harman, 1973, 13.

easy inference. Would then the notion of implicit belief characterise belief in the ordinary sense? Not necessarily. On the one hand, implicit beliefs are presumed to be consequences of explicit beliefs, and in the storage model of belief there is no reason to suppose that even explicit beliefs are accessible to consciousness. Harman is quite clear about this. He says:

> A belief can be explicitly represented in one's mind, written down in Mentalese as it were, without necessarily being available to consciousness.[96]

Harman's example of inaccessibility is a Freudian one, where inaccessibility is explained by repression. But it may also be the case that an agent clearly entertains a belief but refrains from storing it in his belief box to avoid too much clutter. What Stalnaker objects against is the separation between the contents of the belief box and the role it plays, that is, between beliefs and their use.

On the storage model, one may then have an agent to whom his explicit beliefs are not accessible, and who reasons with beliefs that are not stored. This is why the model does not seem fit to describe anything like an ordinary notion of belief. Stalnaker sums up his main objection against this model of belief as follows:

> Ordinary knowledge is a *capacity*, and ordinary belief a *disposition*. (...) But what is the capacity or disposition a capacity or disposition to do? The storage model has nothing to say about this, and so has little promise of clarifying the problem of logical omniscience.[97]

This consideration leads Stalnaker to develop a different approach of belief all together. It is the 'question-answer machine' model that we shall describe once we have had a look at some background considerations about belief representation.

---

[96]Harman, 1973, 14.
[97]Stalnaker, 1999c, 251.

4.3.2   Belief Content and Belief Representation

As we just saw, Stalnaker rejects the sentence storage model of belief because it does not provide answers to some basic problems connected with belief. Distinguishing between explicit and implicit belief does not help us understand why it is that agents sometimes belief that P and yet do not believe that Q even when the contents of the two are identical. Stalnaker concludes that the sentence storage model does not even ask the right questions that would help us to better understand the 'ordinary' notion of belief for which the notion of accessibility of beliefs is crucial.

Having gone through Stalnaker's criticism of the 'received view' of belief representation, we might expect that his next step will be to develop and defend a position that he finds satisfactory. Yet, this is a step Stalnaker refuses to take. His reasons for abstaining from developing an account of belief representation are, as we shall see now, a very important part of his philosophical enterprise.

We have noted on previous occasions that at the core of Stalnaker's enterprise is the idea that speech is a kind of action, and that the way to account for the effect utterances have, is to look at what changes they cause in the epistemic alternatives the audience entertains. We also saw[98] that, on Stalnaker's account, truth or falsity of beliefs does not depend directly on the causal relation between an agent and his environment. The relation between the environment and the way it is represented by belief is one of counterfactual dependency.[99] So for example, a belief that there is a moose here, tends to be caused by a moose. We are not always right, we may mistake a deer for a moose, but being wrong most of the time would prompt us to change our beliefs about moose. Therefore, we can say that usually our beliefs about moose depend on moose.[100]

---

[98]In section 3.3.5.

[99]For example in Stalnaker, 1999a, 229 he says: 'States of organisms,..., carry information when there exists a pattern of counterfactual dependencies between those states and corresponding states of the environment.'

[100]We may still want to ask what sort of dependency on the states of environment is at play here. As I see it, my beliefs about the dodo are a result of my reading about it and looking at old pictures and new computer reconstructions

People express propositions because they hold certain beliefs and want to influence the beliefs of others. Many philosophers who investigate beliefs, knowledge, and language, seem to think that the next important question that should be answered is how we should represent beliefs, and what makes them true in particular cases. Advocates of causal theories , philosophers like Kripke, Kaplan, Lewis, Salmon and Evans, to name just a few, have spent lots of time and energy trying to defend the notions of singular proposition, direct reference, or haecceitism. What these notions have in common is that they are related to the representation and interpretation of certain classes of propositions. In Stalnaker's case, we cannot tell what his position with respect to these notions is because, while he sometimes analyses their problematic consequences, he does not see the question of representation of belief as a part of his undertaking. He says that

> ...language is a device for achieving certain purposes, and
> we should separate, as best we can, questions about what
> language is used to do from questions about the means it
> provides for doing it.[101]

In Stalnaker's view, propositions carry information, and the exchange and gathering of information is what is crucial to our use of language. All we need to assume in order to use propositions in the ways Stalnaker does, is that beliefs, regardless of how they are stored, help the believer to tell apart the circumstances where the content of the belief holds from those where it does not. This is all Stalnaker claims to presuppose. He says that even the advocates of a more fine-grained approach to belief

> ...must admit that propositions, whatever they are, *have*
> truth-conditions, and that representation, whatever else it
> might be, distinguishes between possibilities. Even if there

---

of the life and times of that unfortunate bird. Not much has changed in my physical environment with respect to the dodo since I was born since the bird, called 'dodo'– meaning 'stupid' in Portuguese, because of its child-like trusting attitude has been extinct for centuries.

[101]Stalnaker, 1999b, 2.

> is some kind of fine-grained propositions that is essentially
> involved in representation, and that is useful for describing
> it, we can agree that anything with representational content
> has truth-conditional or informational content, and that we
> can begin by considering its role in the description and ex-
> planation of rational activities, including speech.[102]

This is not intended as an argument against fine-grained approaches
to propositions. It is rather a low-key defence of his own, coarse-
grained approach. The claim he makes here is that even if it were
the case that fine-grained propositions are involved in belief repre-
sentation, a truth-conditional approach would still have a role to
play.

Even if we agree to separate the question of truth-conditional
content from the question of belief representation, one still should
say something about the relation between the two. Stalnaker says
that

> ...the propositions believed,..., are not components or con-
> stituents of a belief state; there need not be an internal rep-
> resentation of some kind corresponding to each proposition
> believed. Instead, propositions believed are properties of a
> belief state.[103]

One wishes Stalnaker were a little more forthcoming on the subject
of the relation between propositions and beliefs. However, any elab-
oration on that subject would almost inevitably lead him to specula-
tions about the way beliefs are represented. He wants to keep focus
on what propositions are good for. One may find our presentation of
Stalnaker's strategy rather surprising: it might seem that Stalnaker
deals with the subject of belief in numerous articles,[104] but on closer
inspection, what he talks about is propositions and informational
content, leaving the relation between belief in an 'ordinary' sense
and informational content sketched only in very broad contours.

---

[102]Stalnaker, 1999b, 4.

[103]Stalnaker, 1999a, 153.

[104]For example, Stalnaker, 1981, Stalnaker, 1987, Stalnaker, 1989, Stalnaker, 1991, Stalnaker, 1999a.

We have just looked at Stalnaker's criticism of the sentence-storage model of belief. One of the points raised against that model was that it had little if anything to say about the 'ordinary' notion of belief. Now we see that Stalnaker himself cautiously avoids that topic as well. His position is coherent, to be sure, but it is with some surprise that we realise that he makes much fewer claims than would at first appear.

One might think that the propositions we express depend on what we believe, and that what we believe is intimately linked with how it is represented, with how it appears to us as agents. If this were the case, a separation between the form and the content of belief, such as Stalnaker proposes, would be indefensible. Stalnaker tries to deflect this line of criticism, saying that

> . . . one might be skeptical about the very idea that content can be separated conceptually from form and means – from the vehicles of speech and thought. Explanations of the semantic properties of expressions in terms of (. . .) the conceptual roles or uses of expressions are examples of accounts that try to avoid contents as objects altogether. The attempt to do semantics without propositional content is motivated more by pessimism about the possibility of an adequate account of propositions than it is by optimism about the possibility of explaining phenomena without them. But I think the pessimism is properly directed against the idea of absolute, context-independent propositional concepts. If we think of propositions as functions from some given domain of relevant alternative possible situations to truth-values, we may be able to reconcile the conceptual distinction between form and content with the phenomena that motivate skepticism about propositions and possible worlds.[105]

Let us see now the model of belief Stalnaker advocates without taking a position on how belief is represented.

---

[105]Stalnaker, 1999b, 3-4.

4.3.3   Question-Answer Machine

The model of belief that Stalnaker proposes[106] is driven by the idea
that belief and knowledge are to be understood as dispositions and
capacities to answer questions.[107] Stalnaker urges us to discard the
metaphor of storing sentences in a belief box, and proposes instead
a hypothetical 'question-answer machine'. There are important dif-
ferences between these two approaches to belief. While the hypo-
thetical question-answer machine needs – among other things – *some*
facility to store information, just *how* it stores information is not rel-
evant to what it knows or believes. All one needs to assume in the
question-answer model, is that if an agent knows or believes that
P, then the information (or misinformation) that P is implicit in
whatever way the information is stored. It is required, though, that
the information that P should be, at least in principle, available to
the agent, meaning that she can access the information in order to
answer a question that calls for the use of P. As we shall see, the
notion of accessibility is pivotal to this proposal.

Accessibility of information is seen as a matter of degree, and
a sharp distinction between explicit and implicit belief is avoided.
Accessibility is also context-dependent. As Stalnaker puts it:

> Attribution of knowledge and belief are obviously highly context-
> dependent, and the line between what we already know and
> what we could come to know if we made the effort may be
> one thing determined somewhat arbitrarily in different ways
> in different situations.[108]

In developing the question-answer model, the main task is to spell
out accessibility conditions for information that is presumed to be
believed. As we just noted, accessibility comes in shades and grades.
One can answer some questions easily, quickly, and with little if any

---

[106]My presentation of it is based on Stalnaker, 1999c, 251-254.

[107]At least at this stage, Stalnaker is mainly interested in 'knowledge that',
that is, propositional knowledge. In Stalnaker, 1999d he briefly mentions the
problems of 'knowledge how' but does not attempt to incorporate it explicitly
into his model.

[108]Stalnaker, 1999c, 252.

hesitation, while coming up with answers to other questions can take considerable time and effort. While the answer to one question may come almost immediately, answering other questions may require numerous steps, and we may even be uncertain at the outset whether we shall arrive at the desired result at all. Just how easy must the search be if the answer is to count as something one already knows or believes rather than something one has the capacity to know or believe? What we know already, what we can derive with various degrees of ease, and what we do not know seems to be a matter of degree. There does not seem to be any natural way of drawing the line between these cases.[109]

What is more problematic in this model is that one proposition can be the answer to different questions, and regardless of what standards of easy access we adopt, that proposition may be easily accessible in response to one question but not in response to another. For example,[110] it would take most people a long time to answer the question 'What are the prime factors of 1591?' but much shorter time to answer the question 'Is it the case that 43 and 37 are the prime factors of 1591?' Yet answers to these two questions have the same content, even on a very fine account of content.[111] Part of the problem is that the questions themselves can change what an agent knows – we all remember the Platonic dialogues where innocent boys are coaxed to give answers to questions they hardly even understood at first. Yet it is a fact, and not a problem, that questions can change what an agent knows. It only becomes a problem if we are trying to use agent's question-answering abilities to get at what the agent could be said to know even if the questions were not asked. On this model, we seem unable to separate knowledge and belief from the use they are put to. Finally, Stalnaker says, "even if we had a satisfactory account of accessibility for the question and answer model, it would not be clear how to generalize it to an account of knowledge and belief in terms of capacities and dispositions to use

---

[109]Stalnaker makes this point in Stalnaker, 1999a, 252, see also the quotation from Stalnaker, 1999a, 253 at the beginning of this paragraph.

[110]This example is given in Stalnaker, 1999c, 253.

[111]In both cases the answer is '43 and 37 are the prime factors of 1591'.

information (or misinformation) to guide not just one's question-answering behavior, but one's rational actions generally."[112]

We have now introduced the basic ideas of Stalnaker's question-answer model. Let us reiterate what they are. Stalnaker rejects the belief-box model of belief because, as he argues, it does not help us to understand the crucial function of beliefs – the way in which agents *use* the information they stored. Turning away from the problem of storing belief, Stalnaker emphasises that having a clearer notion of accessibility of belief is the most important task at hand. Accessibility turns out to be context-dependent and hard to separate from the goals of the agent. Stalnaker's main dissatisfaction with the model he proposed is that having introduced belief and knowledge as dispositions and capacities, it seems very difficult to separate these notions from the particular contexts in which they are used. This problem becomes more urgent if we acknowledge that we may have beliefs that we do not express simply because we never need to.[113] We may have beliefs where it is hard to imagine what they imply in terms of dispositions and capacities,[114] as well as dispositions and capacities that do not translate into speech. Stalnaker treats belief as a disposition, speech as action, and even propositions do not have to be necessarily expressed verbally. Is there nothing special about speech? To those of us who have the suspicion that there is, Stalnaker owes at least a brief sketch of what distinguishes speech from other kinds of action.

4.4   Answering Our Problems

At first glance, and even a second one, I was puzzled as to why Stalnaker thinks that logical omniscience is a problem for his framework. Let me try to explain now where that sense of perplexity came from.

---

[112]Stalnaker, 1999c, 253.

[113]I cannot give an example of such belief because then it would, by definition, become expressed and thus used.

[114]I may for example believe that there is a decimal number expression so long that to read it would take longer than my lifetime.

In this section we shall try to weave together several strands of Stalnaker's work creating a somewhat speculative picture of the connections between his concept of a possible world, his Gricean inclinations, and the problem of logical omniscience. While aiming to remain true to the spirit of his work, some of the claims made here are not based on anything Stalnaker wrote. Some parts of the synthesis developed here are speculations about what Stalnaker should or might have said, rather than an observation of what he actually did say.

4.4.1 Preconditions of Omniscience

It would seem that for a problem of logical omniscience to be a cause of concern to a possible-world framework, a number of preconditions have to be in place. We shall now look at what they are. In doing so, we shall use observations we have gathered in previous chapters – that is why this section has a somewhat recapitulating flavour.

While the problem with logical omniscience arises, strictly speaking, as soon as we characterise belief content in a way that makes it impossible for us to distinguish between two beliefs as long as their contents are logically equivalent, there are several features the presence of which in a framework makes the problem more pressing.

Firstly, the problem of logical omniscience develops in a system where a connection is made between the contents of utterances and the contents of beliefs. We could imagine someone trying to avoid the problem by positing extensional propositions as contents of utterances and structured propositions as the contents of belief. This would be very counterintuitive but the problem would not arise. For the problem to arise, both the contents of beliefs and those of utterances have to be individuated in a suitably coarse-grained fashion. This is the point that seems to trouble Stalnaker the most. He tacitly assumes ordinary modal logic, and explicitly individuates propositions by the truth-values they take in the possible worlds of the relevant domain.

Secondly, the problem can become a source of concern only if we do not change the terms, that is, if we do not change the meaning of

'belief' in such a way that it becomes quite separate from an intuitive notion according to which a person who believes that P is likely to at least assent to P when confronted with it.

Thirdly, the extent to which the problem of logical omniscience develops depends on the kinds of necessity that characterise the individuals, relations, and states of affairs in the domain of interest. For example, if we build into our framework metaphysically necessary properties for individuals and natural kinds, the class of necessarily equivalent statements will be bigger than if we did not take that step. We saw that Kripke's framework abounds in necessities, and because at least some of them are metaphysical, that is, independent of all epistemic considerations, the class of necessary statements in this framework is large and epistemically opaque. On the other hand, if we have a framework which does not enable quantification over sets of possible worlds with more than one member, the notion of necessity we can derive is much weaker. In fact, because in such a framework we cannot express the difference between what is actually the case and what is necessarily the case, identifying the cases of necessarily equivalent statements is difficult to say the least.[115] This we saw to the case with Lewis's framework. We have not yet dealt with Stalnaker's approach to necessity, but it is safe to say that it will turn out to be very different from both Kripke's and Lewis's approach.

Finally, when quantifying over sets of possible worlds, provisions have to be made for quantifying over individuals. This is not really a separate problem, but rather a precondition of quantification over sets of possible worlds. We have seen that Stalnaker refuses to defend any particular position regarding the representation of beliefs, and indeed does not want to say anything about the metaphysics of propositions beyond identifying their information content. Yet since he clearly does quantify over sets of possible worlds, we can try to analyse the preconditions that have to be in place for that quantification to make sense. Presuppositions about identity of individuals

---

[115]It needs to be stressed here that I do not intend my comments to be applied to mathematics. As I said before, considerations regarding mathematics can be, and often are, a separate part of a theorist's framework.

are among them, and they are connected to the question of rigid designation, direct reference, and singular propositions. We can try to infer something about the way in which the notions of rigid designation and direct reference are compatible with Stalnaker's framework without claiming that Stalnaker intends to commit himself to either of these notions. This enterprise is facilitated by the fact that Stalnaker has not been silent on the topic of rigid designation. We saw in section 3.3.6 that he exerts himself considerably trying to account for some epistemically counterintuitive consequences of rigid designation. An analysis of the Mars - Venus problem[116] inspired him to say that in some cases, the meaning of an assertion is to be identified with the diagonal of the propositional concept. On this basis as well as from various examples from his 'Assertion',[117] we may conclude that Stalnaker is open to adopting rigid designation for proper names. However, just what the notion of rigid designation in Stalnaker's framework amounts to depends on the way the domain over which we quantify at a given occasion is constructed. Closely connected with the notion of rigid designation and direct reference is the issue of essentialism and haecceitism. In section 3.2.11, we saw that rigid designation presupposes some version of essentialism and that direct reference is linked to an essentialist form of haecceitism. Stalnaker does not seem to adopt either haecceitism or essentialism, nor, as I shall try to show presently, does he have to choose between the two.[118]

We have now had a little preview of the points one should look at when determining how the problem of logical omniscience expresses itself in a framework. In the following sections we shall have a closer look at Stalnaker's framework from the perspective of these considerations. While we dealt with Stalnaker's work in other sections and chapters, this is the time to a look at the fine print.

---

[116]Stalnaker, 1987.

[117]Stalnaker, 1978.

[118]Putting it this way makes it seem as if Stalnaker could have his cake and eat it. This is no quite the case. We shall try to show that he does not have to make any metaphysical assumptions about individuals in order to preserve the rigidly designating behaviour of proper names.

4.4.2   Possible Worlds Under the Magnifying Glass

In order to interpret Stalnaker's notion of a possible world correctly, one has to keep in mind the extent to which it is context-dependent. He expresses this view in numerous places, but it is the following passage where we find it stated most succinctly:[119]

> I doubt that it is plausible to believe that there is, independent of context, a well-defined domain of absolutely maximally specific possible states of the world, but I do not think the proposed conception of propositional content requires a commitment to such a domain. The alternative possibilities used to define propositions must be exclusive alternatives which are maximally specific, relative to the distinctions that might be made in the context at hand. But one can make sense of this requirement even if there is no ultimate set of possibilities relative to which any possible distinctions might be made. One might think of possible worlds as something like the elements of a partition of a space, rather than as the points of the space. The space might be partitioned differently in different contexts, and there might be no maximally fine partition. (This is only a rough analogy. And the space itself may also vary from context to context.)[120]

In other words, in Stalnaker's framework one cannot use the notion of 'all possible worlds' without further qualification. There is no provision here for an absolute totality of possible worlds – there are always only sets of possible worlds relevant to the situation at hand. We need yet to get a better idea of how the context-given restrictions on sets of possible worlds function, but already we can predict that a contextualisation of the domain of possible worlds has some bearing on the problem of logical omniscience.

Because in Stalnaker's framework belief content is expressed by coarse-grained propositions, a belief in a sentence that expresses a necessarily true proposition still implies belief in all necessarily true

---

[119]Most quotations occurring in this section have been used before, in section 3.3.1. It is in the interest of ease of reference that they be repeated here.

[120]Stalnaker, 1981, 135.

sentences. However, many possible worlds are excluded from the agent's consideration because they are not relevant to the situation she finds herself in. This influences the development of the problem of logical omniscience in two ways: On the one hand, because the number of possible worlds under consideration in a given situation will be relatively small, there will be more equivalences between propositions, which makes the problem worse. On the other hand, the possible worlds under consideration will be epistemically accessible, which makes the problem less pressing. It is important to note, though, that in Stalnaker's framework, many of the propositions that turn out to be equivalent on for example Kripke's framework, will be so only contingently.

The shift from metaphysical possibilities to epistemic possibilities changes somewhat the way the problem of logical omniscience expresses itself in a framework, but does not eliminate it. In Stalnaker's framework, the problem could only be eliminated if the coarse-grained approach to propositions were abandoned, but this, as we know well, is a step Stalnaker will not take.

4.4.3 Shifty Propositions and Unstable Individuals

As we saw earlier, in section 3.3.3, the identity of a proposition depends on the domain over which it is defined. Therefore, once the domain shifts between one context and the next, propositions start shifting with it. As Stalnaker puts it:

> If the alternative possibilities there are vary with the context, then so do the propositions which are, according to the conception of content I am sketching, just ways of distinguishing between the alternative possibilities. One can make sense of questions about identity and difference of the propositions expressed in different utterances or acts of thought only given a common context - a common set of possibilities that the propositions are understood to distinguish between. This yields a conception of proposition, which is less stable than, and very different from, the traditional conception, but

> it is, I think, more adequate to the phenomena of speech and
> thought.[121]

We see now that even though propositions are fully determined by
the truth-values they take in a given domain of possible worlds, there
is a catch to it. In Stalnaker's framework, we cannot talk about *the*
set of possible worlds because each situation creates a new context.
And as the context changes, so does the set of possible worlds that
are relevant to it, as well as the propositions that characterise agents'
attitudes. This point is not new but it is useful to remind ourselves
of it here because it supports our line of argument, which is that
Stalnaker's framework has features that, when taken fully into ac-
count, make the problem of logical omniscience less pervasive that
one might think at first sight.

The changes that propositions undergo between one context and
the next have a bearing on the identity of individuals across possi-
ble worlds. We saw that in Kripke's framework, what is and what
is not possible was, in the end, determined by metaphysical ne-
cessity. Essences, that is, necessary properties, were adopted into
the framework to help to deal with the identity of individuals in a
way independent of any epistemological considerations. Necessary
properties of individuals and of natural kinds facilitated quantifica-
tion over individuals and natural kinds in all possible worlds. They
also helped define the domain of Kripke's possible worlds in virtue
of telling 'possible possible worlds' apart from 'impossible possible
worlds', to use a somewhat awkward idiom. These two functions of
necessary properties were closely linked.

In Stalnaker's framework, on the other hand, the domain of pos-
sible worlds depends on the context and is a function of epistemic
alternatives entertained by agents at a given occasion. As a conse-
quence, there is no inherent need here to provide for the reference of
proper names in all metaphysically possible worlds. It is quite suf-
ficient for proper names to refer only in the domain at hand. Given
that the domain consists of epistemic alternatives, one does not need
to provide a universal guide to the reference of proper names, and

---

[121]Stalnaker, 1981, 135.

one need not to adopt here either haecceitism or essentialism. Rigid designation demands that a designator refers to the same individual in every possible world. If the domain of all possible worlds is specified as a domain of epistemic alternatives relevant to the situation at hand, rigid designation does not have to be accompanied by any metaphysical assumptions, as was the case in Kripke's framework. In Stalnaker's framework, the notion of the *same individual in every possible world* is determined by what the participating agents are happy to consider to be the same individual in a given context, and the question of essential properties is avoided.

We saw that in Kripke's framework, essential properties were an integral part of the system. They were also responsible for generating a class of potentially epistemically opaque necessary statements. In Stalnaker's framework, where metaphysical assumptions are avoided, statements like 'Heat is the motion of molecules' or 'Water is H_2O' are necessary only to the degree to which the agents think they are. The necessity that can be attached to them is epistemic. The class of metaphysically necessary statements has disappeared and what we have in its stead are necessary statements that are *per definitionem* epistemically accessible. Moreover, at any given point we consider only those propositions that are defined over the domain that is contextually given. The class of necessary propositions is profoundly different from the way we thought about it in Kripke's framework, and the difference helps limit the damage the problem of logical omniscience can do.

The obvious objection to be raised against the use of epistemic possible worlds is that there is no guarantee that even if agents agree with each other, their conclusions are not off the mark. They may all agree at some point that 'this is an elm' or that 'the man in the corner is drinking champagne' and yet all be wrong. Stalnaker considers these kinds of scenarios. It motivates his information-theoretic account of content.[122] The main idea behind that approach is that content should be explained in terms of 'counterfactual dependencies that tend to hold, under normal conditions, between thinker's states

---

[122]We have dealt with that in section 3.3.5.

and their environment.'[123] The information theoretic account does not have a direct relevance to the problem of logical omniscience, but should be mentioned here to help us complete the picture of Stalnaker's approach to foundational semantics.

4.4.4   Partial Worlds and Kindred Nightmares

One question that emerges from our considerations of Stalnaker's approach to possible worlds and that is directly relevant to the problem of logical omniscience, is the question of completeness or partiality of possible worlds within his framework. We read in the quotation above that '[t]he alternative possibilities used to define propositions must be exclusive alternatives which are maximally specific, relative to the distinctions that might be made in the context at hand.'[124] This is as far as one can get from Lewisian realism about possible worlds. Stalnaker does not follow Lewis in taking the step of making his worlds as real as the actual world is, and rendering them thereby unsuitable for reasoning about communication. His possible worlds are as specific as the situation requires but not any more than that. With respect to an intuitive concept of an actual world, seen here as the background source of the epistemic alternatives, they could be called partial. But that would be a misnomer. With respect to the situation they describe and the function they are supposed to fill, they are complete. One might say that with respect to, let us say, a discussion about the possible outcome of the war in Iraq, a proposition describing the belief that Manchester United is simply the best is undefined, but when someone expresses, in the middle of a discussion about Iraq, his views about Manchester United, that statement will shift the context of the discussion.

    The context-dependent approach to possible worlds has repercussions for the role they can play in the modelling of linguistic competence. The other lesson of the football example above is that in Stalnaker's framework, we would need a separate account of what it means to master a language. The conversation participants are

---

[123]Stalnaker, 1990, 140.
[124]Stalnaker, 1981, 135.

tacitly assumed to be competent speakers, and Stalnaker does not give an account of what that means.

The situatedness of possible worlds, the way they are specific relative to the situation at hand, changes the scope of logical omniscience dramatically. An agent cannot be said to entertain a belief Q in virtue of entertaining belief P unless both beliefs are equivalent *and* relevant to the situation under consideration at that point. The possibility of beliefs 'running away' with an agent is limited.

None of the above considerations removes the disadvantages of individuating belief contents by truth-conditions completely. We saw, however, that Stalnaker's framework is less vulnerable to the problems of logical omniscience than one would have thought. We shall now look at some of the finer workings of the building of a context. The extent to which Stalnaker can connect his thinking about possible worlds with pragmatic considerations of context-building helps determine the extent to which he can put some flesh on his framework.

4.4.5 A Griceful Interlude

Stalnaker's thinking about context has been inspired by his interest in speech acts. Although Stalnaker does not mention the work of Paul Grice often, its influence on his own work should not be underestimated. Stalnaker mentions Grice's ideas in Stalnaker, 1978, and then comes back to this topic over twenty years later, saying:

> My initial concern was with speech, and my approach was inspired and heavily influenced by the work of Paul Grice in which it was argued that we should see speech as action to be explained, like any other kind of action, in terms of the beliefs and purposes of the agent. Language is a device for achieving certain purposes, and we should separate, as best we can, questions about what language is used to do from questions about the means it provides for doing it.[125]

---

[125]Stalnaker, 1999b, 2.

Grice's view of speech as action had a profound influence on Stalnaker. Seeing communication as a goal-oriented activity is also helpful once we start looking at the ways of describing and restricting context. In Grice's work, the notion of context is treated only implicitly, by describing, so to say, the rules of the game of conversation. Stalnaker learnt Grice's lesson, and used it to describe context in more formal terms. What Stalnaker and Grice share is the emphasis on information as the main vehicle through which language fulfils its function.

Grice[126] suggests that conversation is ruled by certain maxims that can help us to analyse various ways in which conversations can be awkward, inefficient, misleading or strange, ways that have nothing to do with the truth or falsity of what is said. Conversation is seen, in Grice's work, as a purposeful enterprise governed by the *cooperative principle*, which says: 'Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.'[127] The principle divides into four categories of maxims, related to the quantity, quality, relation, and the manner of what is said. Let us look at them now, and then see what use Stalnaker made of them.[128]

The category of quantity delivers the following maxims:[129]

> Make your contribution as informative as is required (for the current purposes of the exchange).

> Do not make your contribution more informative than is required.

Under the category of quality falls a supermaxim saying 'Try to make your contribution one that is true', and two more specific maxims:

---

[126]For example in Grice, 1975.

[127]Grice, 1975, 45.

[128]I assume that most readers are familiar with these maxims: it is for the purpose of comparison that I find it useful to repeat them here.

[129]This and the immediately following overviews of maxims of conversation are taken from Grice, 1975, 45-47.

> Do not say what you believe to be false.

> Do not say that for which you lack adequate evidence.

Under the category of relation falls a single maxim, namely 'Be relevant.' Finally, the category of manner produces the a supermaxim 'Be perspicuous' and various maxims such as

> Avoid obscurity of expression.

> Avoid ambiguity.

> Be brief (avoid unnecessary prolixity).[130]

> Be orderly.

These maxims outline the requirements an utterance should meet in order to function well as an assertion in rather general terms. The generality does not allow us to see what happens when a speaker violates one of these maxims (and such things happen, especially in spoken exchanges, frequently). We know that a competent speaker can, and does, carry out so to speak on-line repairs of what he hears. Stalnaker's approach to the analysis of assertion (which uses propositional concepts) allows us to see in more detail just how things can go wrong and what speakers do to rectify faulty assertions. Stalnaker adapts Gricean maxims to his framework, using the notions of proposition, possible world, and context set in the following manner:

> A proposition asserted is always true in some but not all of the possible worlds in the context set.

> Any assertive utterance should always express a proposition, relative to each possible world in the context set, and that proposition should have a truth value in each possible world in the context set.

---

[130]If having a problem understanding the wording of this particular maxim, I suggest the reader refers back, two lines up, to the first maxim of manner.

> The same proposition is expressed relative to each possible
> world in the context set.[131]

Stalnaker does not claim that every assertion actually meets the
requirements listed above – he says only that if an assertion is to
function properly, it should meet these requirements. Should any
of them be violated, the speaker may still be understood but the
conversational exchange will be less than optimal.

Let us now have a look now at how this definition of assertion
maps onto the Gricean maxims. The first requirement tells us not
to say things that are trivial, that we should contribute informa-
tion. We can relate it back to the category of quantity. The second
requirement relates to the category of relation, telling the speaker
to say things that are relevant. The last item of the definition is
related to the category of manner. The point is that a speaker who
is ambiguous is at danger of being misunderstood. The category
of quality is very important to information exchange. We can re-
late it to the first condition on assertion but only with qualifications
because Stalnaker's description of assertion deals with changes in
beliefs, not with whatever makes beliefs true.

The definition of assertion helps Stalnaker to shift the focus from
describing speech acts to a more structured analysis of information
exchange. The Gricean heritage is still very visible here, but Stal-
naker focuses on the level of changes in the information states of
agents, which are characterised by their propositional content.

In Grice, Stalnaker found an ally in a study of conversation from
the viewpoint of information exchange. The emphasis of his work is
in describing how language is used to change agents' beliefs. That
is why he takes so seriously the danger of epistemically undesirable
consequences of the apparatus he uses. If logical omniscience were
indeed pervasive throughout his framework, it would undermine its
primary goal. However, I hope to have shown that at least on my
somewhat speculative (yet textually supported) interpretation of his
work, Stalnaker should not worry so much.

---

[131]Stalnaker, 1978, 325.

4.4.6 Stalnaker's Limits

We have now traced the problem of logical omniscience through Stalnaker's work. This helped us understand the intended use of possible worlds in Stalnaker's framework as well as some of the less-obvious properties Stalnaker's possible worlds exhibit. We saw that while the problem itself does not disappear, it changes and shifts. In the end, however, the only way in which it could be removed would be by adopting a notion of proposition that is more finely-grained, closer to language. Stalnaker seems well aware of this, and it is a consequence of his views that he would be willing to accept. In his defence of the coarse-grained notion, he points to the strengths of his framework, and reminds us that a lot of the advantages would be lost if a more fine-grained approach to propositions were adopted.

A comparison between Stalnaker and Kripke helped us understand the extent to which the aims of the two authors differ. With its consistent emphasis on the situatedness of communication, the main strength of Stalnaker's framework is in the areas of foundational semantics and semantics of modal statements. With respect to descriptive semantics, he seems to adopt a broadly Kripkean picture, as we could see for example in his attempt to develop an intuitively plausible account of making necessarily false assertions (the Mars - Venus example, in section 3.3.6). Descriptive semantics, however, does not seem to be in the central to his work. If it were, he would investigate more closely the changes that would be brought into a Kripkean descriptive semantics once Stalnaker's possible-world framework is adopted. Stalnaker does not intend to say much about language competence or about belief representation in general. There are theories that he is happy to argue with, as we saw in his criticism of the sentence storage model of belief, but his disagreement with those theories has not prompted him to develop a theory that he would want to defend.

4.5   Conclusion

In this chapter, we focused on the notion of proposition, the conditions that have to be in place if we are to make a good sense of it in a framework, e.g., the notion of a transworld individual and the kind of necessity that characterises a given framework. We saw that in Lewis's framework, propositions have a limited application. This seems to be an inevitable consequence of the absence of any kind of essentialism in that framework. Indeed, it seems that the very notion of worldbound individuals does not capture the way speakers use counterfactual statements, and some kind of essentialist commitment may be necessary if a framework is to capture the way language is used. In Kripke's work, we found propositions in their most characteristic form with all of its disadvantages. The problems seem to stem, at least to a large degree, from the externalism present in Kripke's framework. In Stalnaker's work, the concept of proposition changes and some of the problematic consequences which we saw in Kripke's framework are mitigated. In this part of the chapter, we focused on the connection between the kind of possible-world framework, its essentialist commitments, and the kind of necessity it implies.

In connection with our investigation of the notion of proposition in Kripke's work, we looked at the Pierre puzzle and some of the variety of responses it elicited. In connection with Stalnaker's work, we focused on the problem of believing necessary propositions and concluded that while the situation is in this respect better than in Kripke's framework, the main problem, logical equivalence of necessary propositions, is still present.

In the next chapter, we shall propose an approach to the descriptive semantics of proper names that builds on the lessons learned here.

Part II

Chapter 5
User-friendly Descriptive Semantics

5.1 Introduction

In this chapter, we shall look at what we learned from our analysis of descriptive semantics, both in Chapter 2, which focuses of descriptive semantics, and in other parts of this thesis. We shall use the desiderata implicitly present therein to outline a descriptive semantics that avoids at least some of the drawbacks of the theories we investigated while preserving the insights that seemed valid even under close inspection.

In the course of this work, we have seen time and time again how closely connected are the three aspects of semantic theory – the descriptive and the foundational semantics, and the semantics of modal statements. Oftentimes what may be tolerable within one part of a semantic theory leads to undesirable consequences in another. A good example of this is Kripke's adoption of essential properties, which was helpful in building his possible-world framework, but led to undesirable consequences in what had to be presupposed about speaker's competence (foundational semantics). We shall keep in mind the interconnectedness of the three parts of a semantic theory when outlining our own approach to the descriptive semantics of proper names.

5.2 Why Not Descriptivism?

In this thesis, we chose not to explore the – once very influential – thesis that names are hidden descriptions.[1] What motivated this decision? One could respond that none of the authors whose work we analysed advocates this view. But this answer begs for an explanation of our choice of authors. I chose to analyse the theories I did because they are a representative sample of different influential approaches to the semantics of proper names, as well as of different ways of building and using a possible-world framework. That is also why I focused on the primary literature – there are refined and modified versions of Kripke's, Stalnaker's, Lewis's and Kaplan's theory around but I was interested in analysing the theories in their original form, where the distinct features, the advantages and drawbacks of the basic approach are hopefully easiest do detect. I did not come across a descriptivist theory of proper names whose influence nowadays is comparable to that of the theories I chose.[2] That may be a consequence of seeing the task of a semantic theory of proper names in a particular way: if we expect a theory of proper names to deal with their descriptive and foundational semantics and the semantics of modal statements, then the theories that do not adopt this approach will seem less attractive. Let me try to explain this claim now.

In my view, there are some issues in the descriptivist approach to reference that should be, and often are not, clearly separated. In his *Naming and Necessity*, Kripke says that there are basically two ways of understanding the descriptivist claim:[3] we can see descriptivism as claiming that descriptions give us the *meaning* of a proper name,

---

[1]It may at first sight seem a little confusing that in this section we shall talk about 'descriptivism' as well as about 'descriptive semantics.' However, I think it is quite clear that the similarity of words does not imply a connection of the concepts they name.

[2]I do not mean to slight the contemporary study of descriptivism, where Stephen Neale's book, (Neale, 1990), deserves an especially close attention, and Gary Ostertag's reader (Ostertag, 1998) is an invaluable guide to the contemporary discussion.

[3]Kripke, 1980, 59.

or as making a weaker claim, namely that they *fix its reference*. If the former were the case, then the meaning of a proper name would be identified with a cluster of descriptions (or a weighted majority of descriptions) that are true of the referent of the name in question.[4] If the latter were the case, descriptions would be simply

> ...pedagogical devices employed in teaching the name to someone who does not know how to use it. Once our student has identified the object to which the name applies he can forget or ignore these various descriptions by means of which he identified the object, for they are not part of the sense of the name; the name does not have a *sense*.[5]

The properties mentioned here are said to be commonly attributed to the referent, but they are sometimes also supposed to be constitutive of the referent's identity.[6] As Searle puts it, "To ask for the criteria for applying the name 'Aristotle' is to ask in a formal mode what Aristotle is; it is to ask for a set of identity criteria for the object Aristotle."[7] If we see descriptivism as making the weaker claim, that is, that associated descriptions help speakers identify the referent of proper names, then it makes a claim only about the foundational semantics of names. If that were the case, I would be quite sympathetic to it.[8] But if that were the case, descriptivism would still owe us an analysis of the semantic value of proper names. Moreover, the weak reading is not the one intended by descriptivists, as Searle, who considers both the weak and the strong reading, argues.[9]

---

[4]Kripke claims in Kripke, 1980, 61 that this – making such an identification – is what Strawson does in Strawson, 1959. Searle, in Searle, 1967 seems to be doing the same.

[5]Searle, 1967, 91.

[6]Both characterisations in Searle, 1967, 95.

[7]Searle, 1967, 94.

[8]I think I agree here with Stanley, 1997, 568, who shows that the weak reading of descriptivism is compatible with a variety of positions on the descriptive semantics of names, including a Kripkean one.

[9]In Searle, 1967, 90-92, the 'weak' claim is described as a saying that "explaining the use of a name by citing characteristics of an object is not giving the rules for the name, for the rules contain no descriptive content at all. They

If we see the strong reading as the intended one, then descriptivism posits a very strong relationship between the descriptions and the meaning of a proper name, and that meaning is what we would call the name's semantic value.[10] In its Russellian form, descriptivism analyses away what we would call proper names,[11] claiming that they are, in fact, descriptions. Russell says:

> The names that we commonly use, like 'Socrates', are really abbreviations for descriptions.[12]

and

> . . . the name 'Romulus' is not really a name but a sort of truncated description. It stands for a person who did such-and-such things, who killed Remus, and founded Rome, and so on.[13]

This position generates many problems, one of which I would like to stress in particular: it is generally accepted (by, i.a., Frege, Mill, Kripke, Salmon, Linsky, and Evans) that proper names are a separate kind of expressions.[14] If that is the case, analysing them away,

---

simply correlate the name to the object independently of any descriptions of it." Searle then goes on to characterise the 'strong' reading, saying that "proper names do have a sense necessarily but have a reference only contingently. They begin to look more and more like shorthand and perhaps vague descriptions." According to this reading, proper names "refer only on the condition that one and only one object satisfies their sense." In the passages that follow, Searle clearly identifies the 'strong' reading as the intended one. Moreover, we can also claim with confidence that Russell, though he did not consider the two readings as such, would have, at least in Russell, 1956, chosen the strong reading as the intended one.

[10]See for example Searle's assertion that proper names "have essentially a sense and only contingently a reference – they refer only on condition that one and only one object satisfies their sense."(Searle, 1967, 92).

[11]I am well aware that for a period of time Russell claimed that proper names are expressions like 'this' and 'that', and that what we would call a proper name is a description. I shall, however, follow the common usage of the term 'proper name' even when describing Russell's position.

[12]Russell, 1956, 200.

[13]Russell, 1956, 243.

[14]In chapter 6, I present various kinds of data to support this claim.

i.e., reducing them to another kind of expression, is a move that would require more of a justification than Russell provides.

In the later version of descriptivism, especially in Searle's work,[15] the descriptions, while still constituting the meaning of a name, are not, strictly speaking, identified with it. The problems connected with specifying the relation between a name and the descriptions connected to it led him to say that if one asks

> whether or not proper names are logically connected with characteristics of the object to which they refer, the answer is "yes, in, a loose sort of way."[16]

To sum up: descriptivism does not claim that descriptions only serve to fix the reference of a proper name. In its Russellian form, it posits something close to an identity between the meaning of a proper name and descriptions associated with it, and these are then equated with the names' semantic value. In the more sophisticated versions of descriptivism, for example in Searle's work, this claim is moderated but the result seems to be only a lack of clarity as what precisely the connection between a proper name and its associated descriptions is.[17]

The Russellian and Fregean position,[18] on which the psychological content, the semantic value, and the information potential are all conflated, is not attractive. Part of its lack of allure lies in the fact that while I see a semantic theory as divided into a descriptive and a foundational part depending on the problems it sets out to solve, the Frege-Russell view is rather undifferentiated in this respect.

Searle is aware of the problems that beset the Frege-Russell view and tries to avoid at least some of them by loosening the connection between a proper name and its associated descriptions. But even

---

[15]I have in mind Searle, 1967.

[16]Searle, 1967, 96.

[17]Witness the way Searle formulates his position in the quote we used above, saying that the two are connected in 'a loose sort of way.'

[18]Even though I have used here examples from Russell's work, there are many points of similarity between Frege's and Russell's position. In many contexts the term 'Frege-Russell' position is used. We shall also for our purposes overlook the differences between Frege's and Russell's view.

in his proposal, it is not clear whether he makes claims about the descriptive semantics of proper names, that is, about the systematic contribution that names make to the truth-value of sentences in which they occur, or about their foundational semantics, that is about what it is about the way names are used that makes it the case they have the semantic value that is attributed to them. As long as the connection between proper names and descriptions is unclear, it is hard to see how we would even begin to evaluate his proposal within the framework we have adopted. We have been relying on making use of the distinction between reference and existence, semantic value and speaker's meaning, and properties of language versus the epistemic preconditions of its use. From that perspective, the main descriptivist claim regarding the logical relation between proper names and descriptions either conflates the issues of the descriptive and the foundational semantics or else it underspecifies their relation.

In order to outline a descriptive semantics of proper names that we would find attractive, and that is what we are after here, we should keep in mind the distinctions named above, and see which insights of the theories we analysed we want preserve.

5.3   The Non-essential Kripke

Let us turn our attention now to the currently perhaps most influential theory of descriptive semantics of proper names – Kripke's theory. In analysing it, we did not find a problem with the thesis that names are rigid designators – we had a problem with how this thesis is implemented. Our main objection was directed at the essentialist assumptions that accompany rigid designation in Kripke's framework.[19]

It seems intuitively plausible that proper names are rigid designators if what we mean by that is the following:[20]

---

[19]For some of my objections see section 2.8.

[20]The definition below is cited in Kaplan, 1989b, 569. We have used it before, on p. 47. It comes from Kripke's letter to Kaplan, and its chief attraction is

> (**RD**): A designator d of an object x is rigid, if it designates x with respect to all possible worlds where x exists, and *never designates an object other than x with respect to any possible world.*

Prima facie, if we replace d by a proper name, (**RD**) is a claim about the behaviour of proper names in modal statements. For it to be false, a proper name which designates x in the actual world could designate something other than x in a counterfactual situation. Given some assumptions that I shall spell out a little later, I cannot see that such a situation could arise, and have therefore no objection against (**RD**).

In addressing the descriptive semantics of proper names, Kripke relates his views to those of John Stuart Mill. Mill is supposed to have claimed that names have denotation but no connotation (using more modern terms, the claim is that names have reference but no sense).[21] Elaborating on this claim, Kripke says that while there is a connotation (e.g., associated descriptions) connected with a name, it is not a part of the meaning of a name.[22] The meaning of a name is its reference. In other words,

> **Millianism**: The semantic value of a proper name is its bearer.

Millianism is a claim about the descriptive semantics of proper names, that is, about the systematic contribution proper names make to the truth-value of sentences in which they occur. We can see Millianism as a precondition of rigid designation, and as we had no problem with rigid designation, we find no fault with Millianism either. If rigid designation and Millianism form the hard core of Kripke's doctrine, then it is a core we want to preserve. However, as we shall see now, there are some other claims related to this core of Kripke's position on the descriptive semantics of proper names that are, in our view, more troubling.

---

that it was designed to stand in-between the persistence and the obstinacy of rigid designation of proper names.

[21]Millianism was introduced in section 1.7.3.

[22]Kripke, 1980, 26.

5.4   Where Does Essentialism Belong?

Some of Kripke's best known arguments deal with the idea that some properties are essential. We have analysed and criticised Kripke's essentialist views at length in section 3.2. Our objective here is to look more closely at the role essentialism plays in Kripke's proposal, and see whether we may preserve the claims about the descriptive semantics of proper names we endorsed in the previous section, and remove those that we had found objectionable, that is, his essentialist claims.

We shall start by investigating whether Kripke intendes his essentialism to be a part of the descriptive or the foundational semantics.[23] Someone might object that since Kripke introduces essentialism in order to be able to deal with the semantics of counterfactual statements, the question whether it belongs to the descriptive or the foundational part of a semantic theory of proper names is beside the point. We should, however, bear in mind that an analysis of counterfactual statements is not just a part of a study of possible worlds. As Kripke expresses very succinctly,

> ... we do not begin with the worlds (...) and then ask about criteria of transworld identification; on the contrary, 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.[24]

We can thus view an analysis of counterfactual statements as a useful device that helps to bring out various features of our language that may otherwise be left untreated. It remains to be seen, however, whether those features, and the explanations proposed to account for them, belong to the descriptive or the foundational semantics. Therefore, the question about the place of essentialism in a semantic theory still stands.

---

[23]Kripke's essentialism extends to natural kinds in general. However, right now we shall focus only on those claims that have a direct bearing on an account of semantics of proper names.

[24]Kripke, 1980, 53.

In Kripke's framework, 'essential property' is used as synonymous with 'necessary property', and the notion of necessity relevant here is, according to him "not a notion of epistemology but of metaphysics."[25] This is not new to us – Kripke pays a great deal of attention to the dichotomy between the notions of a priori versus a posteriori, which are epistemological notions, and that of necessary versus contingent, which are metaphysical notions, and we have dealt with this issue at length in section 3.2.5. What we need now, is to establish the nature of the relationship between metaphysics and language within Kripke's theory. Because essentialism is a metaphysical doctrine, it has to do with the existence (or conditions of identity) of individuals to which essences are ascribed. For example, when discussing Elisabeth II, Kripke says: "How could a person originating from different parents, from a totally different sperm and egg, be *this* very woman?...It seems to be that anything coming from a different origin would not be this object."[26] On the other hand, in virtue of assessing the admissibility of some counterfactuals (e.g., 'Elisabeth II could have been the daughter of Mr. and Mrs. Truman.'), essentialism with respect to proper names seems to be connected also with the descriptive semantics.

Given that rigid designation is a property of proper names in the language as we use it, and that necessary properties have a direct bearing on the functioning of rigid designators, we can read the following quotation as supporting our claim that essentialism is intended to be a part of the descriptive semantics of proper names.

> When I say that a designator is rigid, and designates the same thing in all possible worlds, I mean that, as used in *our* language, it stands for that thing, when *we* talk about counterfactual situations. I don't mean, of course, that there mightn't be counterfactual situations in which in the other possible worlds people actually spoke a different language.[27]

If we use *our* language, with *our* meanings, to describe counterfac-

---

[25]Kripke, 1980, 35-36.
[26]Kripke, 1980, 113.
[27]Kripke, 1980, 77.

tual situations, then it would seem that if a particular counterfactual situation cannot arise, for example one where Elisabeth II has different parents from those she in fact has, it is, in Kripke's view, a function of what we mean by the referring term, here 'Elisabeth II', that this is the case. Using an example about a natural kind, Kripke makes a point that supports this claim:

> We could have discovered that that the actual cats we *have* are demons. Once we have discovered, however, that they are *not*, it is a part of their very nature that, when we describe a counterfactual world in which there were such demons around, we must say that the demons would not be cats.[28]

It is useful in this context to remind ourselves of the reasons that inspired Kripke to reject descriptivism. Most of his criticism centres around the observation that the properties that are ascribed to individuals by descriptivists are contingent, and cannot be therefore a part of the meaning of the term if the meaning is to give us any indication about the behaviour of the term in counterfactual situations.[29] Let us now look at the reply Kripke gave to Searle. When Searle says: "I am suggesting that it is a necessary fact that Aristotle has the logical sum, inclusive disjunction, of properties commonly attributed to him,"[30] Kripke replies: "Such a suggestion, if 'necessary' is used in the way I have been using it in this lecture, must clearly be false. (*Unless he's got some very interesting essential property commonly attributed to Aristotle.*)"[31] Searle's claim is that the meaning of a proper name is, in some way, constituted by the descriptions commonly associated with it. Kripke's reply is that this cannot be the case unless some of those descriptions correspond to essential properties. We have seen that essential properties play a role in determining the reference of a term in Kripke's framework

---

[28]Kripke, 1980, 126.

[29]See for example the discussion in Kripke, 1980, regarding Aristotle (p. 60-63), Moses and Jonah (p. 66-67) or the celebrated Gödel-Schmidt example (p. 83-84).

[30]Searle, 1967, 95.

[31]Kripke, 1980, 61, my italics.

(for example, there is no world in which cats are demons given that they are not demons in the actual world). Therefore, even though the folklore has it that for Kripke, the meaning of a proper name is its reference, the essential properties, if there are any, are a part of the meaning. In this way, the essential properties, and Kripke claims to identify some, are a part of the semantic value of a proper name. Consequently, essentialism is a claim about the descriptive semantics of proper names.

Our problem with essentialism, as we have shown in section 3.2.9, was that it relied on scientific realism. We should stress that scientific realism is not a scientific view, or a natural consequence of scientific discoveries. It is a metaphysical doctrine. Now that we have shown that essentialism is a claim about the descriptive semantics of the term in question, it becomes even less palatable. What Kripke seems to claim is that some claims of a particular kind of science (physics, genetics) are, so to say, built into our language. If we take into account our history, and the variety of languages spoken in the world, many of which are used by peoples to whom the concepts of physics are quite foreign, we cannot but try to do without essentialism, at least in the from presented by Kripke.

In order to try and do without essentialism, however, we have to briefly look at the role it is supposed to play because only if we find an adequate replacement, can we say that we can do without it.

5.5 Where Should Essentialism Belong?

We are now in a position to appreciate to role of essentialism in Kripke's work. Let us have a closer look at the main problems connected with essentialism as a part of the descriptive semantics of proper names.[32] When dealing with Kripke's semantics of modal statements,[33] we pointed out two important features of the essentialist strategy. Firstly, we have shown that essentialism cannot be

---

[32] The following objection could be extended to the use of essentialism in the treatment of natural kind terms, but we shall keep our focus on proper names.

[33] In section 3.2.

derived from an analysis of the functioning of proper names. Nathan Salmon came to the same conclusion, and expressed his view in words with which I fully identify (and with a display of eloquence to which I cannot but aspire):

> No conclusive reason has yet been given to suppose that the essentialist principles entailing the necessitations of (the necessary a posteriori) statements are anything but what they appear to be: theses stemming from an irreducibly metaphysical, and philosophically controversial, theory of essentialism, a theory which is no mere consequence of the philosophy of language.[34]

This is why I shall now argue that we should not accept essentialism, as least in this from, into our descriptive semantics of proper names.

Secondly, we have seen that some essentialism was needed for the functioning of the rigid designation of proper names. Essentialism had an important function in defining the domain of possible worlds which Kripke considered. His possible worlds were metaphysically possible worlds, and in determining the truth-value of counterfactual statements, Kripke had the option of using essentialist principles. Therefore, if we want to do without essentialism as a descriptive doctrine, we have to propose something that would fulfil that role.

Before moving any further, we should note that the term 'essentialism' could be used in two distinct ways. It is important to make this distinction in order to be quite clear about what it is that we want to reject or endorse. We shall distinguish a *weak* and a *strong* version of essentialism.

Weak essentialism endorses the claim that there are properties which an individual must have in every possible world of a domain if it is to exist in that world. An individual has such a weakly essential property necessarily with respect to the domain under consideration. For example, if we work within a domain of epistemic possible worlds, and it can be assumed that for every speaker we consider it is unimaginable that the actual individual Charles, the prince of Wales, could be a daffodil, then *not being a daffodil* is

---

[34]Salmon, 1982, 217.

Charles's essential property in that domain. Weak essentialism does not imply that there should be any overriding or metaphysical necessity attached to any essential property.

Strong essentialism is a position according to which one can ascribe an absolute or metaphysical necessity to essential properties. This does not imply a commitment to any particular strongly essential property. For a position to be strongly essentialist, the commitment to the possibility of strongly essential properties is all that is required. Kripke defends a strongly essentialist position, as we see in his treatment of some a posteriori necessary statements. He says that

> ...characteristic theoretical identifications like 'Heat is the motion of molecules', are not contingent truths but necessary truths, and here of course I don't mean just physically necessary, but necessary in the highest degree – whatever that means.[35]

I suggest that essentialism in its weak, non-metaphysical form can be treated as a part of the foundational semantics of proper names. Let us now look at the reasons behind this move, and discuss its consequences.

We have characterised the foundational part of a semantic theory as one that deals with what it is about the use of a particular kind of term that endows it with the semantic value which the descriptive semantics claims it has. Regarding the descriptive semantics of proper names, we have found no good reasons to suppose that names are not rigid designators. We also did not come across any reasons that would lead us to reject the claim that the semantic value of a proper name is its reference. We do therefore concur with Kripke, Kaplan, and others in supposing that names are rigid designators, and that their semantic value is their reference. But unlike Kripke or other proponents of strong essentialism, we argue that the background for rigid designation, that is, the conditions of identity of individuals, must be supplied by the foundational semantics of

---

[35]Kripke, 1980, 99. We have used a part of the same quotation previously, on p. 96.

proper names. Identity conditions of individuals should reflect the way names are actually used, and cannot therefore be supplied by a metaphysical doctrine, which strong essentialism undoubtedly is.

When using a proper name, speakers aim at referring to its bearer. In order to identify the bearer in the actual world, they use a whole variety of criteria, most of which they themselves would see as contingent, i.e., tied to a particular time and place, the type of discourse, and various assumptions they make about the knowledge their audience has of the intended referent. Problems regarding the identity of a referent – that is, questions as to whether a proposed referent of a name still counts as the bearer under certain circumstances – do not arise often in our everyday discourse, but when they do, we do not consult our linguistic knowledge (competence) but rather non-linguistic sources, such as law, history, medicine, or forensic pathology.

When a speaker is faced with a statement involving an individual where her intuitions fail, she will, using her best knowledge, look for the most generally accepted criteria for that kind of statement. For example, when trying to establish the truth-value of a statement that says 'There was a period of time during World War II when Poland ceased to exist', she would look at the historical data, and consult legal theories about successor states.

For most counterfactual statements speakers are likely to encounter, there are generally accepted criteria by which they should arrive at the statement's truth-value. These criteria are a part of what is taken in a given society to be common knowledge. Science has contributed to the content of what is generally accepted, and to that extent it too provides criteria of identity of individuals. In some border cases (e.g., the beginning and the end of a person) the jury is still out.

5.6 What Do We Learn From Counterfactual Statements?

Counterfactual statements serve to sharpen speakers' criteria of what it takes to be a particular individual or a kind of individual. Speakers are, as we pointed out, aware that in everyday life they often use contingent criteria for the identification of individuals, or that they use a mix of criteria some of which may be essential within a given discourse. Just consider what a typical missing person call looks like:

> Police in Kettering are appealing for help to trace a 32-year-old man who has gone missing from the town. Patrick Cash was last seen at his home address in the town at 5.40 p.m. on Wednesday (November 10, 2004). He is described as white, about 5ft 8ins tall, of a slight build with ginger/blond hair, which is in an outgrowing crop style, and speaks with an Irish accent.
>
> He was last seen wearing a black anorak, dark tracksuit bottoms and black trainers. When he left he also took with him a silver/black Mountain Bike. Officers are keen to trace Patricks bike and would like to speak to anyone who may know where the bike is, have been asked to look after the bike or who has bought a similar bike of anyone fitting Patricks description in the past few days.... Anyone who has seen Patrick, or knows of his whereabouts or Patrick himself, are urged to contact Kettering Police on 01536 411411.[36]

It is quite clear that Patrick would still be himself even without his bicycle or anorak, and with a new haircut. But if a person otherwise fitting his description were found to be 5 inches shorter than Patrick, it would probably not the person the police are looking for. For the police, his height, his build, and his accent (as long as he is alive) might be the properties any person who could be identified as

---

[36]This is an actual missing persons report from Northamptonshire Police, England, available at http://www.northants.police.uk/missing.asp, so in the unlikely event you have seen Patrick or his bicycle, please, contact the authorities.

Patrick Cash would need to have. These are his essential properties with respect to the police search. But in another context, Patrick's build may not be essential: one could say that Patrick would not be as thin as he is if he ate more fish and chips as a child.

The point is that once we reject the idea that a domain of possible worlds can be co-determined by strongly essential properties, that is, once we reject metaphysical assumptions, we are left with domains of possible worlds that are context-dependent. It is then within a given context that some properties become weakly essential. It is not the case, as Kripke claims, that

> . . . the properties an object has in every counterfactual world have nothing to do with the properties used to identify it in the actual world.[37]

I want to claim the opposite: We have nothing to go by but the criteria we use to identify objects in the actual world. Some of them are clearly contingent, some of them are not, some of them are a part of folk-psychology, others a product of science. Using the appropriate criteria involves assessing correctly the kind of context at hand, and the relative usefulness and accuracy of different types of criteria.

Kripke says:

> Mathematics is the only case I really know of where [adequate necessary and sufficient conditions for identity] are given even *within* a possible world, to tell the truth. I don't know of such conditions for identity of material objects over time, or for people. Everyone knows what a problem this is. But, let's forget about that.[38]

I suggest we should not forget about it. It seems quite incontrovertible that most people most of the time do very well at identifying objects in both actual and counterfactual situations. A semantic theory should try to explain this fact. Kripke seems to ask the

---

[37]Kripke, 1980, 50.
[38]Kripke, 1980, 43. We used this same quotation before, on p. 90.

wrong sort of question – asking for necessary and sufficient criteria of identity of various kinds of objects outside of any particular context presupposes that the question has a determinate answer, that there is some context-independent fact of the matter as what what makes a table a table or what makes an individual to be that particular individual. We do not make this assumption.

One might object that the view proposed here implies an endorsement of a 'relative identity.' One could say that if in different contexts different kinds of criteria are used to determine whether someone is Patrick Cash, then each context generates a different Patrick Cash. This would be a very undesirable consequence of our view – if it were justified. But let us see why this is not the case: the referent is what it is, and nothing else. We interact with the referents of proper names and indexical expressions in various ways, and they are the source of information we gather about them. Individuals are what they are regardless of our ways of conceiving of them. This is not to say that our ways of conceiving of individuals are completely arbitrary. There are traditions, conventions, and all manners of more or less efficient methods at play when it comes to the way we perceive objects around us, and humans in particular. Research has shown that especially regarding face recognition, there are also innate factors at play.[39] In fact, a deficiency in the face-recognition model is often a hallmark of a developmental disorder, e.g., autism.[40]

In referring to individuals we use language for our needs and purposes, and what is an individual's essential property in one context, may in another context turn out to be its contingent property. Different sets of necessary properties are posited in different contexts, and they play a crucial role in determining the truth-value of the corresponding sets of factual or counterfactual statements made within a given context. Once we clearly separate the question of existence from the question of reference, the problem of 'relative identity' disappears.

---

[39]See for example Bower, 2002, Valentin and Abdi, 2001, and Bednar and Miikkulainen, 2000.

[40]See Carver and Dawson, 2002.

It is all well to say that referents are what they are and no other thing, but it does not get us very far when we want to establish what kind of facts are relevant to establishing the truth-value of particular counterfactual statements. By relegating the properties that co-determine the truth of a particular counterfactual to the realm of foundational semantics, we acknowledge that sometimes the necessary conditions of identity of the referent are determined by a convention. Often, there is no context-independent fact of the matter as to what still is a particular entity x and what is x no longer. Sometimes the question of truth of a particular counterfactual statement is decided by a specific legal convention, legal decision, medical test, scientific test, decision of a medical ethics board, or even just general agreement. Take for example the question of statehood. Names like 'Germany', 'Russia', or 'Slovakia' are proper names, and those countries are legal entities. When questions about their identity arise, they are often relegated to the realm of law. In deciding whether the contemporary Slovak Republic is a successor state of the Slovak State, which existed during World War II, we rely on legal guidelines. There are no hard facts at play here.

I have shown that in our everyday life we apply various kinds of criteria to determine the identity of an individual. Yet some people may feel uneasy about the claim that there is no context-independent matter of fact as to whether an individual, say George, is still or already himself. These issues are rife with all sorts of ethical and religious considerations. But the only answers that make sweeping, across the board, claims are based in metaphysical convictions. Our language usually reflects a certain lack of determination of our views. For example, we say 'John Smith will be buried on March 1' which may be taken to imply that the individual, John Smith, will be buried, yet strictly speaking that would be horrendous. What this illustrates is that depending on the context, there may be different kinds of conventions at play, and sometimes more than one convention can apply but we still usually know which one is intended by a particular speaker in a particular situation.

What notion of context have we been employing here? While the notion of possible world we have been using is essentially Stalnake-

rian, the context is understood more broadly than in Stalnaker's analysis of conversational exchanges. In his analyses, Stalnaker uses a highly simplified notion of context that allows him to monitor the effects of particular utterances on the belief sets of the participants.[41] The notion of context we need here is broader that the one used in building propositional concepts. We intend a notion which characterises a kind of discourse (e.g., scientific, forensic, medical, everyday), and that notion is related to the notion of common ground, which Stalnaker introduces in Stalnaker, 1978, 321.

5.7 Conclusion

In this chapter, we proposed a picture of descriptive semantics motivated by our dissatisfaction with various features of the Lewisian and Kripkean frameworks. Our proposal is admittedly inspired by our reading of Stalnaker's work, though I would not venture to claim that Stalnaker would subscribe to it. Our basic idea was to propose a notion of transworld identity that does not rely on strongly essentialist assumptions, and yet preserves the notion of rigid designation for proper names. In essence, we made the identity criteria for individuals context-dependent.

We have tried to preserve the those features of Kripkean framework that we had found useful and intuitively plausible, mainly the notion of rigid designation. As we saw in the previous course of this work, rigid designation necessitates some degree of essentialist commitment. Even if one rejects Kripke's scientific realism, an account that incorporates rigid designation has to account for the identity of individuals across possible worlds. As we saw in Lewis's proposal, to give up on the notion of transworld individual has wide-reaching and negative repercussions on the ability of such framework to provide a natural account of modal statements involving individuals. One may worry that any notion of transworld individual implies an unattractively strong degree of essentialist commitment. What we tried to do here, was to show that one can have transworld individ-

---

[41]We examined Stalnaker's notion of context in section 3.3.4.

uals and rigid designation without essentialism that amounts to a metaphysical claim. We suggested a notion of weakly essential properties, that is, properties necessary of an entity within a particular context. This, in turn, relied on our notion of possible world, which, we claimed was derivative of Stalnaker's notion (or, perhaps, just our reading of it).

This chapter also gives an implicit answer regarding the relative place of descriptive semantics and modal semantics. In the course of our previous investigations, we had repeatedly pointed to the close connections between the two enterprises. And while one may still opt for treating descriptive semantics separately from semantics of modal statements, we have tried to demonstrate that the two are so closely connected that that should be at least conceived of as two parts of the same undertaking. A view on the descriptive semantics of proper names is tested by the predictions it makes about the behaviour of proper names in modal statements. Ontological commitments (e.g., essentialist ones) which may go unnoticed or remain hidden in a descriptive account of non-modal statements, come to light when that descriptive account is applied to modal statements. And an account of modal statements that could, as such, be quite attractive, can be largely discredited if it fails to make plausible predictions for the semantics of non-modal statements. For these reasons, we opted for a joint account. Ultimately, in both the descriptive and the modal enterprise, both our linguistic and our modal intuitions should be employed in order to get at an attractive result.

Chapter 6
Foundational Semantics: Names, Indexicality, and Ambiguity

In this chapter, we shall try to motivate and develop a foundational semantic of proper names. Our approach shall be motivated by some pragmatic considerations but we shall also use them to motivate a treatment of ontology of proper names, which could be seen as underlying a descriptive semantic treatment of proper names.

At the core of our proposal is a perceived need to account for the fact that a name can stand for multiple bearers. Our account was inspired by an article by Pelczar and Rainsbury,[1] as well as more recent linguistic literature. We shall look in some detail at theories that try to explain a name's ability to refer to multiple bearers by positing indexicality, and compare it with the usually less popular proposal to treat names as ambiguous.

We start by examining some current theories that point to the indexical or demonstrative character of proper names, and try to assess their advantages. In section 6.2.1, we look at some linguistic evidence that can be interpreted as a reason to think that proper names share some important features with indexical expressions. We then look at differences between ambiguity and indexicality, and try to support a claim that names are, in fact, ambiguous, rather that indexical.

---

[1]Pelczar and Rainsbury, 1998.

In section 6.2.2, we briefly review – again – Kripke's view on proper names, and focus on the points where incorporating either indexicality or ambiguity of proper names would be problematic. Having concluded that Kripke's theory should be made more specific regarding its scope, we look for further inspiration in Kaplan's account of ambiguity of proper names. We adapt his proposal of ontology of proper names and try to show why it is much less counterintuitive than a cursory look may suggest.

Having all this at our disposal, we put it to work in dealing with some problems concerning the functioning of proper names in real discourse. We use, once again (in section 6.8), Stalnaker's notion of propositional concept and illustrate its usefulness in dealing with context resolution of proper names. Finally, we look at various kinds of data that are relevant to comparing the ambiguity view I champion with the indexicality view I try to argue against.

6.1   Outlining the Problem

Let us start by giving a preliminary outline of the discussion regarding the indexicality of proper names. This should enable us to clarify which problems we want to treat, and to identify some objections against a position such as we shall try to develop.

6.1.1   Indexicality, Overt and Hidden

When reading the pre-70s literature on the semantics of proper names, one can notice the almost omnipresent idealisation that there is one bearer for each name. The perfectly commonplace fact that a proper name, say 'John' or 'London', can refer to a number of individuals used to be left out of a theory of semantics of proper names. The one-bearer-per-name idealisation has been adopted by Fregeans, descriptivists, and causal theorists, as well as many others working within the analytic tradition.[2]

---

[2]For example, Frege, 1893, re-print in English 1952, Kripke, 1980, Linsky, 1977, and Dummett, 1973b.

Starting with Burge's well-known article,[3] indexicality of proper
names has been discussed, and several explanations of their indexical
behaviour have been proposed. Proper names have been described
as being indexical in several ways.

Firstly, there is the discussion of the so-called 'hidden indexi-
cality', which is usually connected with a Kripkean discussion of
causal chains or chains of communication.[4] Hidden indexicality
was originally described in connection with the reference of natu-
ral kind terms,[5] and only later attributed to proper names as well.
It describes the semantically relevant properties of the 'backward-
looking' connection between an utterance of a proper name and the
origin of that name (the original baptism). It is related to the claim
that proper names display direct reference, the implication being
that speakers do not need to have individuating knowledge of their
intended referent as long as they are connected, in the right way, to
the chain linking them to it. The discussion of hidden indexicality
also aims at describing further conditions under which reference can
succeed (e.g., the role of the referent's essential properties in the
evaluation of counterfactual statements).

Secondly, there is the open or overt indexicality. As I mentioned
above, this discussion started with Burge's article (Burge, 1973), and
has recently been taken up by Segal (2001), Haas-Spohn (1994), Re-
canati (1993), and others. Burge sees himself as defending a version
of the predicate view, continuing a line of reasoning taken by Russell
and Quine. He points out that we need to account for the fact that
a name, say 'John' can refer to numerous individuals, and suggests
that we should treat names as complex demonstratives consisting
of a name-predicate and a demonstrative, thus analysing an occur-
rence of, for example, 'John' as 'that John'. A proper name, 'John',
then denotes all bearers of that name, that is, all referents that were
given the name in an appropriate way. Together, these referents
form the extension of a predicate, while the circumstances of a par-

---

[3]Burge, 1973.

[4]We find this topic treated e.g., Haas-Spohn, 1994, and Lerner and Zimmer-
man, 1984.

[5]The *locus classicus* is Putnam, 1975a.

ticular utterance help resolve the demonstrative, thus determining which John is intended at a particular occasion.

The name-predicate Burge has in mind is not an abbreviated cluster of descriptions, as it was on the descriptivist view. It is also not an abbreviation of an artificial predicate like 'Socratizes', as Quine once proposed.[6] Names do not convey any information except that the bearer of a name 'N' is an entity that can be called 'N'. Names are predicates in their own right, or, strictly speaking, they are general terms, which together with a copula and an indefinite article are parsed as predicates. In its referential use, a name functions as a predicate conjoint with a demonstrative. The demonstrative helps pick out the intended referent. Thus, Burge says, outside of a context, the sentence 'John is 6 feet tall' lacks a truth-value, much like the sentence 'That book is green'. This is the gist of Burge's proposal.

Others, chiefly Kaplan,[7] have rejected the predicate approach. They want to account for the phenomenon of multiple bearers by postulating an *ambiguity* whereby a name gives rise to a potentially unlimited number of phonetically identical yet distinct names that form distinct words. Which name is used at a particular occasion is then determined by the context of utterance.

We have thus on the one hand the hidden indexicality – related to the origin of a name – and on the other hand either overt indexicality or ambiguity, both of which deal with the interpretation of a name in a particular context. We can see both of these discussions – one about the hidden, the other about either overt indexicality or ambiguity – as a response to related but distinct sets of problems, and as two steps in the treatment of the semantics of proper names. When describing the behaviour of proper names, we want to know how hearers interpret proper names in the context of a discourse. This task has at least two aspects. On one hand, we need to investigate the relation of proper names to their bearers. This relation helps us determine the truth or falsity of statements containing proper

---

[6]Quine, 1961a, 140-143.

[7]In Kaplan, 1990.

names, and is thus a part of descriptive semantics of names. We need to try and spell out this relation in terms that are, at least in principle, accessible to the speaker so that she can meaningfully aim at making true statements using proper names – we need to make plausible foundational semantic claims. On the other hand, we need to determine how speakers and hearers arrive in their production and interpretation at the intended referents. The problem we have to deal with at this level is that proper names seem to have multiple bearers. Our task is to describe the process of interpretation (or perhaps disambiguation) of an utterance of a proper name. In order to describe the behaviour of proper names exhaustively, we might have to investigate both the hidden and the overt indexicality, that is, both the relation of a proper name to its referent, and its relation to the interpreter. Throughout this whole project, however, we have avoided a discussion of causal chains, and though that might deserve a further justification, we shall continue to do so. Instead, we shall present a view of how names are attached to their bearers in our proposal of ontology of proper names (in section 6.3.1). That is why we shall not discuss the hidden indexicality of proper names.

We shall focus on the second step, that is, the process of interpretation or disambiguation of proper names. There is a perception (e.g., in Haas-Spohn, 1994, and Recanati, 1993) that overt indexicality and ambiguity are competing explanations of the same phenomenon. We shall look at this claim in more detail, and see what the crucial differences between the two positions are. Our aim is to present a version of the ambiguity view, point out its advantages, and answer some of the objections mounted against it.

6.1.2   From The Indexical Point of View

Françqois Recanati[8] has presented a well-articulated version of the indexical view. It is more elaborate than Burge's somewhat sketchy proposal on which it builds. Recanati wants to defend a number of theses: a) that proper names are directly referential; b) that there

---

[8]Recanati, 1993, 140-143.

is a convention associating a name with its referent, and c) that a proper name 'N' means 'the entity called N'. It seems plausible that we could read a) as amounting to a claim that the meaning of a proper name is its referent. Thesis c) is an offspring of the Quinean approach to names, which, as we saw, was taken up by Tyler Burge. At the core of Recanati's approach is the claim that proper names do not have a meaning in the sense of a conceptual content. They are rather like predicates in the sense that a name 'N' can be truly predicated of an entity x just in case the x satisfies certain conditions – a proper name 'N' can truly be applied only to an entity *called* 'N'. The convention that makes this happen (that makes it the case that an entity may truly be called 'N') is, according to Recanati, of a social kind, and forms a part of the extralinguistic context. On the other hand, that there are such conventions seems to be a linguistic fact, a part of what defines the category of proper names, and awareness of the existence of such conventions is a part of linguistic competence.

The linguistic convention associated with proper names can be spelled out as follows: "For each proper name there exists in principle a social convention linking that name to a definite individual, called its bearer. This individual is the referent of the name."[9] This general convention then refers to a specific social convention, which associates a particular proper name with its bearer.

> . . . a specific social convention is involved for each proper name (and sometimes more than one convention, as when a name has more than one bearer). Each time a proper name is used, the linguistic convention is appealed to, and a social convention is thereby invoked, viz. a convention linking the name to some definite individual.
>
> In this framework, a proper name refers by linguistic conventions to whoever (or whatever) happens to be the bearer of that name; but who (what) is the bearer of the name is a contextual, non-linguistic matter, a matter of social conven-

---

[9]Recanati, 1993, 139.

tion. The reference of the name thus depends on a contextual factor, as the reference of an indexical expression does.[10]

There are, Recanati admits, some differences between proper names and indexicals. Ordinary indexicals are 'token-reflexive', meaning that the referring relation holds between an expression token and its referent. In the case of proper names, however, reference is determined by a conventional relation that holds between a proper name as a *type* and its referent. "A token of 'N' refers to an entity, which, in the context of utterance, has the property of bearing the name type 'N'."[11] If two entities bear the same name, the reference of the name will vary between one token and another, but the two entities share the same name type.

One may feel that Recanati fails to answer several relevant questions. The indexicality he describes is on a different level than the one involved in the resolution of a proper name in a context where the name may have a number of bearers. According to Recanati, the indexicality of a proper name (even one that only has one bearer) consists in a systematic reference to a social convention which is seen as part of the extralinguistic context. In fact, if one saw knowledge of the conventions connecting proper names to their bearers as a part of linguistic competence (arguing perhaps from the division of linguistic labour and the view of language as overlapping idiolects), proper names would not be indexical on Recanati's view at all. He acknowledges that more that one name-convention may be associated with a name type, but says that "these conventions are not appealed to in the same contexts of utterance, and this is why the reference may vary from one token to the next."[12] We shall show that more that one name-convention (associated with the same name type) can be operative within the same context, and try to explain what guides successful interpretation in such cases.

Recanati says, "the reference of 'N' is the entity which is called N in the context of the utterance."[13] This does not account for

---

[10]Recanati, 1993, 139-140.
[11]Recanati, 1993, 143, notation adjusted for consistency.
[12]Recanati, 1993, 141.
[13]Recanati, 1993, 141, notation adjusted for consistency.

the fact that a name 'N' can name a particular entity x even in abstraction of a particular context. In particular, that x can be called 'N' is not only a result of a successful use of that name in reference to x (that could be just a lucky coincidence), but also a consequence of a practice to use 'N' to refer to x.

If, following Recanati, we think of proper names as types, then we cannot but say that a token used at a particular occasion refers to the entity that bears the type. That is the only way in which Recanati can preserve direct reference in the sense that the meaning of a proper name is its bearer. If we think of proper names as types, we cannot say that a name refers to its bearer in abstraction of a context because then the meaning of that name would be all of its bearers. And that claim is not compatible with the direct reference thesis (and the singular nature of proper names) which Recanati wants to preserve. Moreover, Recanati does not describe the mechanism responsible for a name's reference to a particular bearer in a context. He aims at providing descriptive semantics of names that could account for multiple bearers but as a foundational semantics his proposal falls short.

6.2 Names and Context Dependence

In this and following sections (6.2.1 to 6.7), we shall prepare the ground for our own approach to the problem of multiple reference, which is a version of the ambiguity approach. We start by motivating the need for an explanation of the multiple reference of proper names that should supplement the 'standard' descriptive semantic theories of proper names.

6.2.1 Variability of Proper Names

A proper name, for example 'John', can refer to different entities in different situations. In a given context, however, we usually know to whom a name refers. Various linguistic and non-linguistic features of a situation determine the context of interpretation. Outside a

context (or in an imperfectly determined one), we may be unable to determine a name's reference: when I overhear some strangers talking about John, I am not likely to know whom they talk about. If, however, I overhear someone talking about Julius Caesar, I can assume she is talking about *the* Julius Caesar. This may be seen as a context-sensitivity of names, a feature which names share with indexicals and demonstratives.[14] In order to interpret an utterance of a proper name, one needs to know the context in which it was uttered, and, in particular, the relevant parameters of the context. This is a hallmark of a context-sensitive kind of expression.

Reference of singular terms changes at different rates depending on the changes of the parameter the term depends on. For example, some indexicals sensitive to the time of utterance ('now', 'yesterday') are highly variable, while other indexicals, even if dependent on the same parameter of the context, are more stable ('yesteryear', 'this month'). Varying rate of change is a characteristic which proper names – the context-dependence of which we are yet to describe – share with indexicals. The name 'Anna', allegedly the most common female name in the world, is much more variable than 'Franklin Delano Roosevelt' is. Evans brought to our attention that 'Madagascar' showed a change of variability over time.[15] Nowadays, unlike four hundred years ago, there is usually no need to specify whether we use that name to refer to a part of the African continent or to the island east of it: other things being equal, we refer to the island.

While some names (like 'Julius Caesar') are so stable they might almost seem context-independent, other names can never be properly interpreted outside a context. Names of days are like this. When you hear, 'Mary said she'll come on Monday', you will not know when to expect her unless you know when this sentence was uttered. Names for family members behave in a somewhat similar fashion. Common nouns such as 'mom' and 'dad' show similarity

---

[14]Indexicals are context-dependent expressions. Demonstratives usually require a physical indication, e.g., pointing (deixis). However, because deixis can naturally be seen as a part of context, just like time, place, etc., I shall view demonstratives as a subclass of indexical expressions.

[15]Evans, 1973.

with both the highly variable pronouns like 'you' and proper names like 'John Smith'. In some situations, the common noun 'mom' can be capitalised, and can start behaving like a proper name, so that when speaking for example to my sister, I can say 'I spoke to Mom yesterday.'[16] These expressions stand right in between deictic expressions and proper names.

We set out to find some linguistic evidence for a claim that proper names are context-dependent. The behaviour of various family-relation terms, names of places, and even people's names points to just that. Does this mean that names *are* indexicals? Not necessarily: it seems that while all indexicals and demonstratives are context-dependent, not all context-dependent expressions are what we would usually think of as indexical. For example, how we interpret the term 'bank' depends on both the linguistic and the non-linguistic context, and this context determines whether the speaker is understood as referring to a monetary institution or a shore of a river. Yet, 'bank' is a prototypical ambiguous expression. In short, context dependence does not distinguish between indexicality, where one word can refer to different entities, and ambiguity, where the context helps determine which of the different homonymous words is used.

6.2.2 Problems for Rigid Designation?

In section 6.1.1, I claimed that some older (but still very influential) theories of proper names suffer from not accounting for the fact that names have multiple bearers. As a fair representative of such theories, I chose Kripke's views.[17]

As we have repeatedly pointed out (for example in section 1.7.3), Kripke claims that it is common sense that "the referent of 'N' is N, where 'N' is replaceable by any name."[18] We have also seen that according to Kripke, descriptions, while they may help us find the

---

[16]Both the 'Mom' example and the 'Monday' example come from Pelczar and Rainsbury, 1998, 298-299.

[17]Mainly as presented in Kripke, 1980.

[18]Kripke, 1980, 25, footnote 3, notation altered for consistency.

referent in the actual world, they do not become a part of what is said, the proposition. The semantic value of a proper name in every possible world is just its bearer – i.e., names are rigid designators. Furthermore, we have, in our own proposal (in the previous chapter, see 236) defended the view that rigid designation, one without metaphysical commitments, is indeed a feature of the behaviour of proper names in natural language.

On the other hand, we have already noted that in natural language proper names have multiple bearers. We shall want to try to reconcile the claim that the meaning of a proper name is just its bearer, and presumably just one bearer, with the observation that a proper name has potentially many bearers. If we want to preserve rigid designation, we need to say more about what a proper name is. Giving the individuation criteria of proper names is a precondition for accounting for their pragmatics.

6.3   Introducing Names Into Language

In this section, we shall investigate the ontology of proper names to with the view of reconciling the above-mentioned tension between the claim that names are rigid designators, and the fact that they have multiple bearers.[19] Wedo not intend to give an exhaustive overview of the ontology of proper names – what we are interested in is a proposal that is capable of explaining the above-mentioned tension. In the latter part of this section, we shall turn our attention to some non-standard uses of names because Burge uses those kinds of cases to support his proposal. If, therefore, we can do at least as well as he does, it can count as a point in favour of the ontology we propose.

---

[19]The following account is inspired by Kaplan's ideas presented in 'Words' (Kaplan, 1990), but it is not intended to be an accurate account of Kaplan's views.

6.3.1 What's In a Name?

When thinking of names we can have, depending on our goals, several things in mind.[20] For example, when deciding on a name for a baby, we can go to a bookshop and buy a book about names, a book of the 'How to Name Your Baby' kind. We expect to find in it names common in our culture together with some explanations about their origin and meaning. Names found in such books do not, *per se*, refer. They are a product of tradition. We shall call them *protonames*.

Some protonames are found in many cultures – we can view the names $\mathcal{J}ohn$, $\mathcal{J}ohan$, $\mathcal{J}an$, $\mathcal{I}annis$, and $\mathcal{S}ean$ as instantiations of the same protoname, **John** – while other protonames are fairly language specific. In some traditions, the institution of protonames is not very important and parents aim at original names. However, people's imagination is limited and so even in China, where there is a strong preference for innovative names, one ends up meeting people who have the same given name. In other countries, on the other hand, parents are allowed to give only names that appear in an official list.[21]

We can think of protonames as capable of producing a variety of *names*. Thus $\mathcal{J}ohn$ and $\mathcal{S}ean$, while being instantiations of the same protoname, are two distinct names. There are many names around. Some names help form the repertoire of given names, which are usually chosen by parents. Other names are used to name families, and their application is usually determined by traditions such as passing the family name of at least one parent to the offspring. In the case of other types of names, such as names of villages and towns, mountains, and dogs and cats with pedigree, traditions determining their use are even more varied.

The basic function of names is that though they are non-referring on their own account, they can be used to introduce into language proper names in their referring function. This can be done in two

---

[20]Right now I focus on people's names but we shall see that the account I present is applicable to all proper names.

[21]This is, to a large degree, true of France and of most Slavic countries.

closely related ways. Following Kaplan,[22] we can distinguish between naming and dubbing. Naming is a conventional way of introducing a name into linguistic circulation. To carry out a naming, a special setting is usually required. 'Hereby I baptise you by the name 'Henry'' as pronounced by a priest can be an example of naming. Dubbings are less conventionalised than namings and do not require a special setting. Calling 'Hi, Fatty!' at someone can be an instance of a dubbing. We may thus see a naming as a dubbing that requires some kind of conventional setting. What dubbings and namings share is the resulting attachment of a name to a bearer. We can think of namings as a kind of dubbing but not vice versa. We shall therefore use the term *dubbing* to encompass both namings and dubbings.

Dubbings introduce into the language what we shall call *referring names*. A referring name consists of a name and a particular dubbing, i.e., a unique event of attaching the name to its bearer. There can be many referring names associated with one and the same name. George Bush, the current US president, and George Bush, the former US president, were dubbed using the same name but because their referring names originated in different acts of dubbing, their referring names are distinct.

We have now arrived at a picture which distinguishes protonames – non-referring entities usually handed down within a culture by a tradition; names, which are so to say their offspring; and referring names, always referring to a particular person, which are a result of 'attaching' a name to a particular bearer by an act of dubbing.[23]

To keep track of the notions we have introduced in this section throughout the rest of the chapter, the reader may consult the following overview.

---

[22]Kaplan, 1989b.

[23]In this context, it should be mentioned that I think the semantic of names of fictional entities should be treated separately, and hope to do that in a foreseeable future. I shall try to argue that names of fictional entities are so unlike (normal, mainstream) proper names in their modal behaviour that we cannot expect an account of semantics of proper names to account for fictional names as well. In this way, I would hope to avoid some of the problems connected with non-referring proper names.

Kind of entity	Referring or not	Examples
Protoname, $\mathbf{N}$	Non-referring	What the names $\mathcal{J}ohn$, $\mathcal{S}ean$, and $\mathcal{J}an$ share
Name, $\mathcal{N}$	Non-referring	$\mathcal{J}ohn$, $\mathcal{S}ean$, and $\mathcal{J}an$ are distinct names
Referring name 'N'	Referring, consisting of a name and a dubbing	'John' as referring to John, that is 'N' as referring to x

6.3.2 Non-standard Uses of Names

One may object that we are multiplying entities beyond necessity but before getting too worried about that let us have a look at some pleasant consequences of thinking about proper names in this way – our approach allows us to account for a number of uses of proper names that do not easily yield to rephrasing on other approaches.

Seeing names as predicates allowed for a unified account of some non-standard uses of proper names. Tyler Burge saw it as a part of the motivation of his proposal.[24] We shall show that our present ontology accounts for these cases at least as well as Burge's account. Consider these sentences:

(1) The transaction was traced to a Ralph Esteban.

(2) John thinks he is an Einstein.

(3) There are frightfully many Annas in this world.

(4) The name 'Seán' is commonly spelled in various ways, for example as 'Shaun', 'Shawn','Sean' and, of course, 'Seán'.

In the first sentence, we know the subject of our suspicions only by his name, $\mathcal{R}alphEsteban$, not his or her referring name. This allows for the possibility that it was a group of people who carried

---

[24]Burge, 1973, 429.

out this transaction, or that the name refers to a fictional person invented just for protection from prosecution. The indefinite article draws our attention to the fact that the author of this sentence is not in position to specify a referring name. Burge would interpret the indefinite here simply as an existential quantifier with the extension of the name-predicate, thereby missing both the fictional case and the case of a group of people assuming an alias.

My ontology does not improve on the standard interpretation of the second sentence. We still have to assume that *being an Einstein* refers to a concept, formed of some of the characteristic properties of the most salient Einstein (that is, of *the* Albert Einstein). On the other hand, I do not fare worse than Tyler Burge, who postulates that names are *irreducible* predicates, devoid of descriptive content.

Sentence (3) states that the name $\mathcal{A}nna$ has generated a remarkable number of referring names, i.e., that there are many people called 'Anna'. Burge could paraphrase this sentence as being about the extension of the predicate *Anna*.

The last sentence, (4) is about the hypothetical protoname from which all of the names $\mathcal{S}haun$, $\mathcal{S}e\acute{a}n$, $\mathcal{S}ean$, $\mathcal{S}hawn$ originated. Burge does not seem to be able to express that those names have anything in common.

We have seen now that the ontology proposed in the preceding section allows us to account in an intuitively plausible manner for a number of non-standard uses of names. Part of the reason for this exercise was to show that one does not have to accept Burge's proposal in order to account for cases like those presented above. We saw that using the ontology we propose, we can account even for some cases Burge's framework cannot deal with.

6.4 Calling People by Their Names

When people refer to me by my given name, 'Anna', or even when they use one of the many diminutives commonly associated with my name, they all use the same referring name. The identity of a referring name is not based just on its phonetic form (after all, each

person pronounces it somewhat differently) or spelling. We should think of a referring name, 'N', as determined by a name, $\mathcal{N}$, and a particular act of dubbing, D, that is, 'N' $= \langle \mathcal{N}, d \rangle$. A referring name, '$N$', can undergo series of changes in the course of its use (e.g., when someone is baptised 'William', and is then called 'Willy' or 'Bill'). These variations do not affect the identity of the referring name as long as they are recognised as staying within the limits of convention associated with that referring name (a person may commonly be called by an idiosyncratic version of her or his name) or the name component of it.[25] More specifically, as long as the speaker uses a form of the referring name 'N' that is commonly associated with $\mathcal{N}$ or the individual x who is called 'N', in her community, she usually manages to refer to x using 'N'.[26] For example, if someone is introduced to me as 'Katherine', I can stay on the conservative side and call her just 'Katherine' but I can also try to be a little less formal and call her 'Kathy', because that is a common form of that name. It may turn out, however, that this particular Katherine's favourite form of her referring name is 'Kate', and that is what her friends call her. Using the appropriate referring name, however, does not, by itself, guarantee a successful communication. For that to happen, the speaker needs to meet further conditions, which we shall examine shortly.

When referring to people (or towns or dogs or what have you), we are usually not acquainted with the particulars of the dubbing ceremony by which their name was bestowed on them – nor need we be. The importance of dubbing is normative: a dubbing, d, makes it the case that 'N' refers to x (that is, for example, that 'John' refers to John). This does not tell us, however, how referring names are interpreted on a pragmatic level.

In order to account for the interpretation of proper names, we

---

[25]There are local conventions that outline the amount of phonological and orthographic variation that is likely to be tolerated by hearers. There are also local conventions associated with the way both a name and a referring name acquire diminutive forms.

[26]It is certainly also possible for a speaker who makes his intentions clear to start a new referring chain by dubbing the person x as 'M'.

introduce a new parameter as a constituent of the context of use - a function that ascribes to referring name a degree of salience, *sal*. An expression $\mathcal{N}$ as used in a context refers to the bearer of the most salient referring name 'N' in that context, where 'N' $= \langle \mathcal{N}, d \rangle$.

A hearer hears an utterance containing an expression $\mathcal{N}$, which seems to be intended as a referring name. Without the help of context, he may unable to tell which of the potentially many homonymous referring names the speaker intends. He interprets the expression correctly if he interprets it as expressing a proposition referring to the bearer of the most salient dubbing d of the name $\mathcal{N}$ within the context. The speaker, meanwhile, when making his utterance, intends – at least in the range of cases we now consider – to refer to a particular person. That is why she chooses a suitable referring name 'N' and makes sure (or makes it the case) that it is the most salient dubbing of the name $\mathcal{N}$ in the context. Let us illustrate this by a simple example. There are two friends, Tim and Tom. Tim's sister's name is Jane, and that also happens to be the name of his girlfriend. They talk about sibling rivalry. When Tim then says 'Jane and I always tried to get out of each other's way when the competition was too close', Tom will correctly interpret him as talking about his sister Jane because her referring name is more salient than that of Tim's girlfriend. If, on the other hand, Tim's phone rings, and he reports later to Tom that it was Jane who called, he has to add whether it was his sister or girlfriend because without that, Tom is not in position to tell whom Tim intends to refer to. Tim may be intending to refer to his girlfriend Jane but without increasing the salience of her referring name his intention will be frustrated, and he will not manage to get his point across.

Interpretation is successful if the hearer interprets the name, in the given context, as firstly, determining a particular referring name, and secondly, if that is the same dubbing of a name as the one intended by the speaker. This is how referring names as used by speakers provide us with a standard against which we evaluate the success of the hearer's interpretation. The speaker, in her turn, uses a name correctly if not only she makes sure that the referring name she uses is the most salient dubbing of that particular name in the

context, but also her intended referent was indeed dubbed by the name she uses to refer to it (that is, she cannot perform a new dubbing without ensuring that it is clear that that is what she does and that the intended audience can tell whom she intends to dub).

How can we square this picture with rigid designation? If we think of proper names as those entities which were the object of Kripke's characterisation of rigid designation, 'N' refers to x, then Kripke's 'proper names' should be identified with our referring names. This is not surprising because only referring names refer (unlike protonames and names). So while we could not accept Kripke's claim that 'John' refers to John, because in his framework one could say that 'John' obviously can refer to many people, once we adopt the ontology proposed here, we can say that 'John', the referring name, does indeed refer to John, the person thus dubbed. In this way, we can defend the direct reference claim while accounting for a pragmatically important phenomenon of multiple bearers.[27]

6.5   Individuating Names

We have looked into the context-dependence of proper names, and proposed an ontology that accounts for rigid designation while allowing for the explanation of some relevant pragmatic phenomena. We individuate referring names in a context by the name used and the dubbing. This implies that Tim's sister Jane and his girlfriend Jane have distinct referring names because they were dubbed by distinct acts of dubbing. Furthermore, we have identified referring names as being the suitable candidates for filling the role of rigid designators. A very similar position has been outlined by Kaplan some time ago.[28] It has become known as the 'ambiguity view', and its author formulated it very succinctly as follows:

> . . . two utterances of 'Aristotle' in different contexts may have different contents. I am inclined to attribute this difference to

---

[27]Of course, now we can say with more precision that it is names that have multiple bearers, while referring names have only one.

[28]In Kaplan, 1989a, and Kaplan, 1990.

> the fact that distinct homonymous words were uttered rather
> than to a context-sensitivity of a single word 'Aristotle'. Un-
> like indexicals like 'I' proper names really are ambiguous.[29]

One of the tasks the ambiguity approach faces is to give identity conditions of distinct homonymous names. This applies to my approach as well – we may want to say more than that 'N' $= \langle \mathcal{N}, d \rangle$. As I mentioned earlier, the individuating dubbing may be epistemically inaccessible to speakers.

Ulrike Haas-Spohn (1994) discusses the individuation of names at length, and raises a number of objections against the ambiguity view. Let us have a look at them. It is clear, Haas-Spohn says,[30] that different senses of a phonetic form NF[31] can arise because the uses of NF occurring in the actual world may belong to different causal-intentional networks, and each such network constitutes a different name of the form NF. More generally, distinct names are individuated by distinct usages, U.[32] One could therefore disambiguate names by indexing them with the usages in which they are involved.

The question then arises how the distinct usages themselves are to be individuated. Haas's solution – parallel to her solution of reference to natural kinds – is that name-experts are responsible for individuation. The crucial problem for the ambiguity approach is to specify when an utterance of the name form NF belongs to a particular usage U of NF, that is, when is it an utterance of the name NF_U.

In my opinion, we can meet Haas's challenge if we clearly distinguish a number of issues. Distinct usages, or causal chains, were proposed to provide a foundational semantics, or a speaker-oriented justification, to a Kripkean descriptive semantics. More specifically, Kripke claims that name 'John' applies to John because he was

---

[29]Kaplan, 1989a, 562.

[30]Haas-Spohn, 1994, 136.

[31]I have altered Haas's notation to avoid confusion with my usage.

[32]In the present context, we can think of both Haas's 'causal-intentional network' and her 'usage' as being roughly equivalent to the familiar Kripkean 'causal chain'.

dubbed using the name 'John'.[33] Kripke then suggests that we can posit a hypothetical chain linking each usage of the name 'John' with that dubbing event. Such a chain constitutes a usage in Haas's sense.

One may object, though, that while there may be some plausibility to positing usages in Haas's sense, it is unclear what function they are supposed to perform. If they are meant to be truth-warrants, and be a part of the descriptive semantics of proper names, that is, to play a role in justifying the use of a name to refer to particular person, then they seem superfluous. This is because we do not have to refer to a usage in order to see whether a name was used correctly – all we need to do is to find out whether the intended referent was indeed dubbed using a particular name. If usages are supposed to be a part of the foundational semantics of names, then they seem superfluous again, because on the pragmatic level of interpretation they do not help the audience decide which referring name was intended by the speaker. I suggest that the interpretation of names as used in a context is guided by principles such as I proposed.

Haas-Spohn further points out that if a speaker were to take two usages of NF to be one, she may fail to utter a proper name all together. A similar problem would arise if an entire community were mistaken in taking certain usages of a name form NF to be about one and the same individual, whereas in fact there would be two individuals involved who are being permanently confused.[34] Haas-Spohn concludes that in such cases, the utterances of the name NF do not belong to any usage whatsoever and the sentences containing these utterances fail to express propositions, and have no truth-values. She concludes that this consequence does not appear unwelcome.

I think there is an interesting difference between the two cases Haas considers. In the first case, the speaker can be perceived as not expressing a proposition, which would lead to a halt in communication. In the second case, however, the reference of such a

---

[33]Using our terminology, we can say that the claim is that the referring name 'John' applies to John because he was dubbed using the name $\mathcal{J}ohn$.

[34]Haas-Spohn, 1994, 138.

name can function on the communal level because everyone in the community shares the same misconception. We would need to go to look for different criteria of expressing a proposition in order to spell out what is happening. This will become clearer once we have introduced our notion of context.

Haas's main criticism of the ambiguity approach is this:

> ...if proper names are ambiguous (...), then problematic cases do not involve error or uncertainty about properties of the context, but error or uncertainty about the words that have been uttered, i.e., about form, not content.[35]

This indeed seems to be an unavoidable consequence of the ambiguity view. However, biting this bullet may not be as difficult as Haas-Spohn would make it seem. Her problem is that a hearer who misinterprets a speaker's utterance containing a proper name fails to understand which words were uttered. That, however, just seems to be a different way of saying that the hearer failed to apprehend the proposition the speaker wanted to communicate. It may happen because the speaker's and the hearer's perception of the salience of a particular dubbing in the given context do not match. In other words, as Kaplan says, "semantics cannot tell us what expression was uttered or what language it was uttered in. This is a pre-semantic task."[36]

6.6 Context

In order to determine which referring name was uttered, the hearer needs to know the relevant features of the context. She needs to know which of a number of potentially competing referring names is the most salient one. That helps her to interpret the speaker's utterance in the way it was intended. The speaker, for her part, needs to know which referring name the hearer is likely to consider

---

[35]Haas-Spohn, 1994, 139.
[36]Kaplan, 1989a, 559.

most salient, and adjust the context accordingly.[37]

In general, hearers depend for identification of the intended referring name on antecedent speech acts, Gricean maxims, knowledge of the surroundings, and other pragmatic factors. Speakers have to take all this into account. The interplay between the speaker and the hearer is crucial to successful communication and use of proper names. To capture that interplay, we need a flexible notion of context. For this purpose, we shall make use of our favourite notion of context, that is our interpretation of Stalnaker's notion.[38] Through the notion of presupposition, it takes into account both the knowledge of the surroundings and the antecedent speech acts, and via restrictions placed on an assertion, it incorporates some Gricean maxims. It was not developed for the purposes I intend to use it for but it is an attractive way of representing what happens in communication. And in the examples I introduce below, that is what I intend to do: analyse the effect an utterance containing a proper name has on the hearers.

As we have seen before (in section 3.3.4), each conversation participant is said to have his own context set, but it is a part of the definition of presupposition that each participant assumes that others presuppose everything he or she presupposes. If the participants' context sets match, they give rise to a non-defective context. In the case of proper names, a particular referring name can be presupposed to be most salient in the context. For example, when I speak of Tony Blair, I usually do not have to make sure that my audience knows about whom I talk. There is one bearer of the name who is most likely to be the salient intended referent (i.e., the prime minister of Great Britain), and the context is, with respect to that referring name, usually non-defective.[39]

---

[37]Pelczar and Rainsbury (Pelczar and Rainsbury, 1998) do not try to apply their ideas to the behaviour of names in contexts. What follows is my extension of ideas that originated with them.

[38]Based on Stalnaker, 1978, and Stalnaker, 1981. We presented it mainly in section 3.3.4.

[39]When a person working in argumentation theory speaks of Tony Blair, typically he does not have to specify which one he means either. The context

6.7 Dialogue and Norms

As a matter of fact, context sets of participants of a dialogue do not always match and, as a consequence, a dialogue does not always proceed the way it is supposed to. We saw previously, in section 4.4.5, that in order to characterise how a dialogue should work, Stalnaker proposes a few rules that have to be observed if a dialogue is to fulfil its purpose of manipulating the audience's beliefs in a predictable way.[40] We shall now investigate some problematic cases involving proper names, and see whether we can provide an explanation. Some problems arise when the speaker's and the audience's perception of a salience of a particular referring name differs,[41] but though this may be a common cause of complications connected with the use of proper names, it is not the only one. As usual, things can go wrong in innumerably many ways. The examples presented below hopefully offer at least an interesting sample.

---

is determinate enough to support reference to Tony Blair, the argumentation theorist.

[40]Let me repeat them here for easier reading. In Stalnaker, 1978, 325, it is said that: 1. A proposition asserted is always true in some but not all of the possible worlds in the context set. 2. Any assertive utterance should always express a proposition, relative to each possible world in the context set. 3. A proposition should have a truth-value in each possible world in the context set. 3. The same proposition is expressed relative to each possible world in the context set.

[41]Henk Zeevat once asked me in a personal conversation what would happen if both the speaker and the audience were wrong about the salience of a particular referring name. My answer was that such a situation does not make sense – salience of a name is defined in such a relation to the context that whenever all sides agree a particular referring name is the most salient one in the context, it is so. On this level, there is no higher authority than the agreement of all participants.

6.8   Application to Problems

More than one referring name is salient to the same degree at the same time

Imagine a conversation about politics in October, 2000, in which someone says 'George Bush has a record of being a conservative politician.' The speaker has the presidential candidate in mind. If it were not previously established which George Bush we talk about, confusion might arise. This situation may run two distinct courses:

a) The hearers do not know how to evaluate this utterance because they do not know which proposition it expresses. They are aware of the presence of two competing dubbings of the name $\mathcal{G}eorge$ $\mathcal{B}ush$, and do not want to commit themselves to interpreting the utterance as expressing either of the two propositions it might express. They feel that their choice would be unsupported by facts, and therefore arbitrary. This is why they perceive the speaker as violating Grice's principle 'Avoid ambiguity.'[42] By being ambiguous, the speaker's assertion also violates Stalnaker's second principle, which states that an utterance should express a proposition that has a truth-value in every world of the context set.

Conscious of their predicament, the hearers demand that the speaker increases the salience of one of the relevant referring names. Their reaction is to ask something like 'Which George Bush are you talking about?' A question like this prompts the speaker to provide information that increases the salience of the referring name she intended (by saying, e.g., 'I am talking about the younger one, the one who's running for president.'), which then leads to a disambiguation of her statement.

b) The hearers do not perceive the speaker as ambiguous, and proceed to interpret the utterance 'George Bush has a record of being a conservative politician.' Let's say that the speaker, Agnes, thinks that George Bush Jr. is a conservative, but his father is a liberal. One hearer, Bob, knows her views and agrees with her, but thinks that she talks about Bush Senior, while another hearer, Carol,

---

[42]Grice, 1975, 46.

thinks Agnes talks about George Bush Jr. but thinks that George Bush Jr. is a liberal and his father is a conservative. We can capture this situation by the following propositional concept:

	A	B	C
A	T	T	F
B	F	F	T
C	T	T	F

Within this table, A stands for Agnes, B for Bob, and C for Carol. In the rows, we see propositions as they are understood by the participants, evaluated in the respective context worlds of the participants. In the columns, we see how the context is understood by each of the participants. The diagonal (squares *A-A*, *B-B*, etc.) then represents whether the participant in question sees – within the context as he or she understands it – the sentence as true or false. In this particular case, the rows tell us that A and C interpret the context in the same way (C knows that A talks about Bush Jr.), while B interprets the sentence differently. In the columns, we see that A and B share the same opinions.

This interpretation violates Stalnaker's third principle, which says that the same proposition should be expressed with respect to every world of the context set. A situation where a different proposition is expressed in different worlds of the context points to a misunderstanding having arisen between the participants. This is different from a disagreement about facts, this is a case of understanding the utterance in question in different ways.

A repair of this kind of situation would be guided by broader pragmatic principles: As we said, Bob knows Agnes' views and he further assumes that she is a rational person who does not change her views abruptly. That might give him a reason to double-check his interpretation, by asking Agnes, for example, "Are you really talking about Bush Senior?" The resulting propositional concept would then reflect the sameness of interpretation between all participants as well as their difference of opinion.

A particular dubbing of a protoname has an overriding salience

Imagine I say: 'I got this skirt from Julia Roberts. Not *the* Julia Roberts, my friend Julia.' Now, if I did not add the second sentence, my statement would be misleading. If you heard only the first sentence, you could well doubt its truth. If, however, you hear the rest of what I have to say, you will have no good reason to doubt the truth of my claim. Had I said only the first sentence, I would have expressed the following propositional concept.

	I	U
I	T	T
U	F	F

I stands for the speaker, U stands for the hearer. We see from the propositional concept that while I and U agree on the facts, and I know I did not get a skirt from *the* Julia Roberts, and you know that there's no reason why I could not have gotten a skirt from a friend, we differ in our interpretation of what was said.

Had I uttered only the first sentence, I would have violated Grice's maxim of relation 'Be relevant'. This maxim warrants that people should not change topic without warning. In this case, I would not have changed the topic explicitly, but I would have disregarded my knowledge of the common ground where the most salient bearer of the name $\mathcal{J}uliaRoberts$ is the American actress. By using the name $\mathcal{J}uliaRoberts$ to express a referring name with a very low salience, I would have violated Grice's maxim in effect. This would have led to a violation of Stalnaker's third constraint.

In this example the speaker respects and correctly perceives the common ground – so much is apparent from adding the second sentence. But we can think of a whole variety of cases where the speaker either unintentionally or intentionally does not comply with the common ground. A study of such cases can lead to a clearer view of what it takes to know a proper name.

**A shift of reference occurs without knowledge or intention
on the part of the speaker**

We can try to use the apparatus I have just proposed to analyse
Donnellan's example,[43] which Stalnaker discusses at length in his
1993 article.

> At a party, a man is introduced to a student as the famous
> philosopher, J.L. Aston-Martin. The student had heard of
> Aston-Martin, and knew, before being introduced to this
> man, something of his work. He talks at length with the man
> at the party, and they become long term acquaintances. The
> student continues to believe that the man he was introduced
> to is the famous philosopher, but in fact he is a different
> person with, we may suppose, the same name. ...Donnellan
> suggests that when the student says, on the day after the
> party, 'Last night I met J.L. Aston-Martin, and talked to
> him for almost an hour', he refers unambiguously to the fa-
> mous philosopher and so says something false.[44]

Contrary to what Donnellan suggests in his further treatment of this
case (see below), I think the student says something false not be-
cause his representation of Aston-Martin is dependent on the famous
philosopher, but because in the world as he takes it, the man at the
party and the famous philosopher are one and the same person.

When the student says on the following morning 'I met Aston-
Martin last night' to a friend who was also at the party, their ex-
change can be represented as follows:

	S	F
S	T	F
F	T	F

S stands for the student, F for his friend. We see that although
they interpret the utterance in the same way, they disagree on the
facts.

---

[43]Donnellan, 1970.

[44]Stalnaker, 1993, 310.

This propositional concept shows that the friend knows that the student did not meet the famous philosopher, and that both in the student's and his friend's world, the referring name of the famous philosopher has higher salience than the referring name of the man the student met at the party. If the student were informed at this point that the man he spoke to was not the famous philosopher, he would withdraw his statement, or at least, this is what we think would happen.[45]

> Later, as the student's representation comes to be more richly dependent on facts about the man at the party, it becomes possible to shift to a context in which these dependencies are normal ones, and the information that derived from the philosopher are the distortions. Which context is relevant to the interpretation of the student's utterances of the name 'Aston-Martin' will, Donnellan says, depend on the point of the utterance, the relevant alternatives the student is trying to distinguish between.[46]

What proposition the student expresses by saying (*): 'Last night I went with Aston-Martin to see the Yankees', as he gets to know the man better, depends not only on the student's intentions, but also on the shift in salience of the name 'Aston-Martin' within his circle of friends. We can suppose that the salience of the referring name of Aston-Martin, the student's friend, will gradually increase over time until it becomes more salient than the referring name of Aston-Martin, the famous philosopher. At this point, when the student utters (*) he says something true, and that is not only because of his intentions, but also because of how he is understood.

---

[45]When we try to interpret the propositional concept above, we can see the limits of this method of representation: given our story, we can infer what the student thinks is the case, and we also know what his friend thinks is the case. We do not know, however, and that is a general point, what would the speaker (or, for that sake, the hearer) say if he or she were in a situation they consider non-actual. That is, both speaker's and hearer's beliefs are sometimes less defined with respect to counterfactuals than what the propositional concept presents them to be. The propositions as understood by agents (on the horizontals) sometimes express more than the agents themselves know.

[46]Stalnaker, 1993, 310.

Stalnaker's example is similar to Evans's Madagascar example.[47] In both cases, the common ground is made defective through ignorance. Evans treats the Madagascar example by claiming that the island becomes the referent of the name $\mathcal{M}adagascar$ when, and only when, the island becomes the dominant source of information about the intended referent. In my view, what is responsible for the shift of reference is the increase of salience of that dubbing using the name $\mathcal{M}adagascar$, according to which this name refers to the island.

Layers of reference

I want to suggest a view on shift of reference of proper names, which, while still rather sketchy, seems to yield intuitively plausible results. Shifts of reference are a result of an interplay between several kinds of reference. What is involved here is the *speaker's reference*, which is the reference the speaker uses to guide her production. Speaker's reference picks out the referent the speaker intends to refer to using a referring name. Secondly, there is the *communal reference*, which picks out the object that is perceived to be the salient referent of a name within a group or community. This reflects the hearer's side of things. We should, however, imagine the audience to be extended to the whole relevant community. We could analyse the communal reference in counterfactual terms: the most salient referent (i.e., the bearer of the most salient dubbing of the name $\mathcal{N}$) is what most members of a community would pick out were they a party to a particular use of that protoname. If I thus start speaking of Julia Roberts, then unless I make further provisions, I will be taken to refer to the American actress because within my community most people, were they presented with the name $\mathcal{J}uliaRoberts$, would take it to refer to the American actress.

The third kind of reference we need to consider is the *semantic reference*, which characterises what a referring name refers to in abstraction of its use. Semantic reference is grounded in the bestowing

---

[47]Evans, 1973, 195-203.

of a name on a person by the act of the dubbing[48] that is, 'N' = $\langle \mathcal{N}, d \rangle$. Unlike the speaker's or communal reference, semantic reference is in principle not epistemically transparent. This – as we shall see later – is why in practice it functions usually only as a kind of default condition.

If everything proceeds as smoothly as it should, the three kinds of reference coincide and if they do not, speakers take it into account (this is the Julia Roberts case). Sometimes, though, by design or by ignorance, they split up and separate.

In Evans's example, Marco Polo mistakenly identified the referent of the most salient referring name 'Madagascar' to be the island off the African coast. He wrote about Madagascar in his travel diaries, and his speaker's reference led to the establishment of a communal reference in Europe according to which the name 'Madagascar' referred to the island. This eventually resulted in a new semantic reference, a new dubbing – it became all right by all semantic lights to use 'Madagascar' to refer to the island. This sort of example forces us to broaden somewhat our understanding of 'dubbing'. So far, we have understood it as a particular act somewhat modelled on a name-giving ceremony for a child. This is not really applicable in this case. Here, we have to understand it as a hypothetical point at which it became correct to use $\mathcal{M}adagascar$ to refer to the island.

Stalnaker's example is a little less complicated: the student's reference to Aston Martin has over time built up a communal reference within his group of friends, where the most salient bearer of the name 'Aston Martin' became the student's friend, and not the famous philosopher. Semantic reference remained unchanged because the name $\mathcal{A}ston\mathcal{M}artin$ existed as a component of the referring names of the two men right from the start.

---

[48]Which may then be said to start chains or trees of use, but as we argued these are not really an essential part of either the descriptive or the foundational semantics of proper names.

Some referring names may turn out to be referent-less

The following problem is presented to us by the history of science:

> On March 26, 1859, a French medical doctor and amateur astronomer named Lescarbault claimed to have observed a planet closer to the sun than Mercury; he called it Vulcan. He calculated the planet's movements and sent the information onto Jean Le Verrier, France's most famous astronomer.
>
> Le Verrier had already noticed that Mercury had deviated from its orbit. A gravitational pull from Vulcan would fit in nicely with what he was looking for. Le Verrier checked other reports and found that other astronomers had also seen a small black disc against the background of the sun. From his calculations, Le Verrier came to the conclusion that Vulcan was 13 million miles from the sun and that it took twenty days to circle that star.
>
> Over the next few years, others reported seeing Vulcan and textbooks added the new member to their lists of planets. But there was controversy because some astronomers couldn't find Vulcan. Le Verrier explained it away by saying that most of the time the planet would be lost in the sun's glare. He said that the best time to observe Vulcan would be during a solar eclipse. The next eclipse would be on March 22, 1877.
>
> Many astronomers had their eyes focused on the sun that day but no one could find the elusive planet. One year later two American astronomers observing a solar eclipse from separate places in Wyoming and Colorado claimed to have seen the lost planet. And they were the last persons to have seen Vulcan. If the planet really exists, no one can find it.[49]

The conclusion is that nowadays we do not believe that Vulcan exists. But how should we treat the reference of 'Vulcan'?

I suggest the following: Mr. Lescarbault, thinking he saw a planet, thought he had performed a dubbing. His letter to Le Verrier, a highly respected astronomer, led to a spread of the refer-

---

[49]Author unknown, was to be found in January 2003 at http://members.tripod.com/TonySakalauskas/index-2.html.

ring name 'Vulcan' within the scientific community. Lescarbault's speaker's reference resulted in the establishment of a communal reference. After 1887, it was discovered that the dubbing performed by Lescarbault was invalid because the ostensive element of the dubbing had failed. The semantic reference had therefore also failed because regardless of what he thought he saw, Lescarbault could not have seen the planet Vulcan. No one could – the planet does not exist. The semantic reference (based on the dubbing) eventually led to a failure of both the speaker's and the communal reference.

What is interesting is that when Le Verrier was communicating with his colleagues about Vulcan, both the speaker's and the communal reference were in order. Writing to his friend, 'Vulcan takes 20 days to circle the sun', Le Verrier expressed the following propositional concept.

	LV	FA
LV	T	T
FA	T	T

LV stands for Le Verrier, FA for his friend, a famous astronomer.

The communication was successful. Everyone in the astronomical community of pre-1877 would have agreed. What went wrong was the dubbing, and, as a consequence, the semantic reference.

This example illustrates that even though usually all we need to take into account is the speaker's and the communal reference, there are cases where we have to turn to the semantic reference. When dealing with failure of reference, we can see one of the limits of the way we represent communicative situations: nothing in the propositional concept is inherently tied to reality. This is a consequence of using Stalnakerian epistemic possible worlds. As long as reference on the communal level functions smoothly, there is nothing in our representation to warn us of a deeper failure. This is intuitively plausible as long as what we aim at is a representation of beliefs of relevant agents (which was a part of Stalnaker's motivation). A failure of semantic reference is not a part of this picture as long as the agents involved are not aware of it.

How does the reference to the planet Vulcan function today? I suggest that we can only speak of Vulcan in nonveridical contexts. Analysis of that sort of contexts, however, falls outside the scope of the present investigation.

Astronomers thought Vulcan existed. When it was proven not to exist, it more or less fell into oblivion. This does not always happen. Sometimes there is an entity that is shown to be a conglomerate of a number of people. This probably holds of Jonah, the prophet, and certainly is true of Bourbaki, a group of French logicians and mathematicians who authored a number of textbooks. In these cases even if we know that there is no particular person to whom the name refers, and semantic reference should fail, the communal reference is still largely successful.

Additional problems with private dubbings

The following example highlights the background of what it takes to use a form of a referring name that is appropriate in the context: Imagine a man, say Mr. Smith, who calls his wife Frances, 'Honey'. Many people have pet names for their partners. Now imagine that a somewhat slow-witted visitor comes to dinner to their house, and hearing the man speaking to his wife comes to the conclusion that her name is Honey. And not just that, he starts calling her 'Honey', too.

Our intuition is that the man is wrong just as I would be wrong if I started calling my friend's mother 'Mom'. But what exactly is going on?

Imagine that you too were present when the visitor later told Mr. Smith, 'Honey made an excellent dinner last night.' Did he say something true? If we were to construct a propositional concept of what the visitor said, we would find out that none of the constraints and maxims we have so far considered were violated. This contradicts our intuition that something did go wrong. In order to account for it in a propositional concept, we introduce a new sort of constraint. This new constraint is based on social conventions regulating the use of names.

It seems that if someone is to know a particular name, he or she must not only know that a particular name can be used to refer to a particular person, but also which form of the intended referent's referring name is appropriate in what context. It is an inextricable part of our ability to use names that we know whether to use a particular form of a referring name just in the family circle, in a circle of friends, or quite generally. We can capture this insight in the following informal constraint: 'Use a particular form of a referring name only if it is appropriate in the context.'

The knowledge needed to comply with this constraint is not strictly linguistic. I could be fluent in English and yet not know how to properly address someone who was introduced to me as 'Cardinal Smith' or 'my husband Jimmy'. A violation of this constraint can lead to misinterpretations.[50]

People occasionally violate this constraint. Sometimes they do it unintentionally – as illustrated by the example we just discussed, sometimes intentionally – name-dropping is an example of that. A person who mentions a celebrity by her/his first name insinuates a degree of intimacy with that person thus seeking to enhance his own social status.

6.9   Ambiguity versus Indexicality

In the three subsections below, I present some data that I see as relevant to the discussion between the two competing theories of the mechanism underlying the phenomenon of proper names' having multiple bearers. The problem we have been facing is that both the ambiguity and the indexicality theory can be modified so that they become rather close in their predictions. An example of this is when we start indexing proper names applied to distinct bearers (e.g., $John_1$, $John_2$, etc.) in the indexicality theory.[51] Nonetheless,

---

[50]That is, give rise to false beliefs: imagine a woman who sits in a café with a male friend and his mother. If both of them call the friend's mother 'Mom', a third person will probably conclude that the male friend is either that woman's brother or her husband.

[51]Haas-Spohn, 1994, Chapter 4, 155.

I believe there is a core difference between the ambiguity and the indexicality approach (and that the ambiguity approach is to be preferred). The following is an attempt to prize the two theories apart, and, taking clues from different areas of research, to show that some data support one theory rather than the other.

Translating Proper Names

As we saw above,[52] Ulrike Haas-Spohn introduced an argument in favour of the ambiguity view that compares the way the indexicality and the ambiguity theory deal with translations of proper names. She says:

> ...the German sentence 'Wilhelm war ein aussergewőhnlicher Mensch' could be translated as 'William was an extraordinary person' if it is about William the Conqueror, but not if it is about Wilhelm Busch; then only 'Wilhelm was an extraordinary person' would be a correct translation.[53]

This points to the fact that a person's name 'N_1' can form a distinct chain of use that is different from another person's name 'N_2' even if those names are phonetically identical. A part of a tradition associated with a proper name can be the way it translates into other languages. This is true not only of person's names, but also of names of places, so that for example Prague, in the Czech Republic, translates into German as 'Prag', but Prague, Ohio, does not. If 'N_1' and 'N_2' were the same name, as the indexicality theory predicts, this phenomenon would be more difficult to explain. An advocate of the indexicality view would have to claim that names with different subscripts should be treated, at least sometimes, as distinct names. However, this modification seems to compromise one of the core claims of the indexicality view – the claim that all people called 'N' share the same name, $\mathcal{N}$. On the other hand, we have to admit that the idea of adding subscripts has been introduced into the

---

[52]On p. 270.

[53]Haas-Spohn, 1994, 135.

indexicality theory independently of potential problems with translations, and it does the job of accounting for translation of proper names.

Determiners Resurfacing?

The indexicality view claims that names should be analysed as consisting of a determiner and a predicate. The determiner Burge has in mind is 'that'. On the other hand, in some languages, like Greek, Italian, Basque, and German, proper names are, either sometimes or often, used with a definite article. This has been viewed as supporting Burge's theory and deserves closer examination.

Native speakers of the above-mentioned languages do not think of this construction (that is, for example of 'il Giovanni') as indicating that names are predicates. Indeed, they seem to be surprised by this proposal. As a matter of fact, there is a number of languages where a construction of a proper name with a definite article can occur, and we shall see that an explanation of these cases does not constitute a straightforward support of Burge's proposal. In German, for example, proper names can appear either with a definite article or without it, thus allowing for 'Ich habe mit Hans/dem Hans gesprochen.' German speakers perceive the difference as purely one of style (the variant with a definite being a less 'correct' one), again indicating that they do not see the article as functioning as a genuine determiner.

We can actually find examples of proper names commonly used with a definite article in English as well. Just consider 'the Nile', 'the Thames', 'the Ukraine', or 'The Titanic' and 'The Hague'. Most of these constructions can be seen as contracted from the likes of 'the river, Thames', 'the ship, Titanic'. They are puzzling semi-idiomatic constructions, and for explanation of their presence one usually has to delve into the history of their usage. But regardless of what the proper analysis of these constructions is, English speakers certainly do not see them as containing a normal definite article, so that 'Nile' in 'the Nile' would be a predicate or a common noun.[54]

---

[54]Some of these examples are suggested in Segal, 2001.

However, we do not even have to go to atypical constructions to start doubting Burge's approach. Examples from spoken English also point to the optionality and special function of a determiner in those constructions where it appears thus casting doubt on the claim that a demonstrative should be seen as an integral part of a proper name construction. Consider:[55]

1a) I live in London.

1b) I live in the London (that is in England).

1c) I live in that London.

2a) Mother is trying to get in.

2b) The mother is trying to get in.

2c) That mother is trying to get in.

3a) This is the/that John (I mentioned yesterday).

3b) *This is John I mentioned yesterday.

The sentence 1a) can be a perfectly idiomatic response to the question 'Where do you live?' posed in most ordinary contexts. In the same context, for example if the addressee is visiting New York, 1b) or 1c) would sound weird.

Sentence 1b) can be used without the added clause with a suitable intonation ('I live in *the* London') where other Londons (e.g., London, Ontario) are not out of the picture. There, the determiner functions in the same way as in the Julia Roberts example (on p. 276), emphasising that the referring names with the highest salience in non-specialised context is intended. In sentence 1b) with the added clause, the determiner provides an anaphoric connection to that clause. In 1c), the 'that' functions as a genuine demonstrative–this sentence can be used when pointing at a map.

---

[55]For examples 1a-d) and 3a,b) I am indebted to Gabriel Segal, Segal, 2001, 560-1.

Example 2a) and 2b) are slightly different. Imagine sitting in a restaurant with your sibling, waiting for your mother. Across your sibling's shoulder, you watch the entrance, and utter 2a). The natural interpretation is that it is your mother struggling with the door.[56] In 2b), the definite article has again an anaphoric function. It could refer to the mother of someone in your company (for example if you are having lunch with a friend and a host of his relatives). And if you utter 2c), you are probably reporting the sight of some woman with a babe in her arms who is trying to negotiate the entrance.

We can easily think of situations where 2a), 2b), and 2c) have each a different reference. The difference between 3a) and 3b) highlights the fact that at least sometimes omitting the (anaphoric) determiner or demonstrative can result in an ungrammatical sentence.

The examples above have shown clearly that a proper name functions differently when prefaced with a determiner. There are rules that regulate when a determiner can be used and when it cannot.[57] Burge's theory does not address this issue. Adding or losing a determiner can be responsible for a change of referent or result in a sentence that is hard to interpret. Burge does not account for this kind of evidence, which leaves his theory in trouble.

6.9.1   Learning from Aphasia

There is yet another area from which we can gather clues relevant to deciding whether we prefer the ambiguity or the indexical theory. The data I am about to present, concern the way language, and proper names in particular, are stored in our brains. My aim is to introduce circumstantial evidence to support the claim that proper names are a category of their own, and thus imply that they do not have a hidden structure of a determiner and a predicate (or a common noun) as Burge suggests. This undertaking only makes sense if we believe that there is a connection between the way our brains function and the way we speak, and that is an assumption I am willing to make.

---

[56]Of course, it could also be your mother-in-law.

[57]I discuss some of these rules a little further on.

Aphasia is a specific language impairment that can be caused by injury or lesion in the language zone.[58] The impairment is often modality-specific, that is, it may concern specifically the understanding or the production of spoken or written word.

In some cases, aphasia affects only a specific class or cluster of terms, e.g., only names of animate entities or only verbs of movement. One certainly would not want to claim that every specific semantically defined class of terms that can be affected by aphasia also constitutes a syntactic category. Fortunately, the fact that proper names can be specifically impaired is not the only fact that can be introduced in the present context, although it is the first issue we are going to introduce.

Of a particular interest to us are cases where while proper names are affected, other parts of the language function remain intact.[59] These cases fall under the description of so-called 'anomia'. Although anomia concerning proper names has been described and studied only for the last 15 years or so, we can find it mentioned in medical literature going as far back as the 15th century.[60] It is therefore clearly not just a byproduct of the interest in proper names within the philosophy of language.

In a clear-cut case of proper names anomia, such as presented by Semenza and Zettin,[61] the patients' command of personal and geographical names was severely disturbed, while common nouns were unaffected, even difficult ones. The patients could, however, retrieve a name when it was presented in a common noun context. For example, in answer to the question 'What sort of bird lives in the San Marco Square in Venice?' they would say 'colombo' (pigeon), but they could not retrieve the name 'Colombo' (Columbus) when asked who discovered America.

---

[58]It can also be a symptom of a degenerative disease, such as Alzheimer's, but these cases are usually more complex and involve a whole scale of symptoms. That is why I will try to exclude the.

[59]Cases are also reported where this happened the other way around – proper names were the only part of language production left.

[60]For example in Guainerio, 1481, cited in Benton and Joynt, 1960.

[61]Cited in Semenza, 1997, 121.

In some cases of proper name anomia, geographical names are relatively less affected than names of persons.[62] Like in other cases of anomia, predicates and common nouns in these anomia patients are intact. The interesting and surprising feature of these cases is that the more readily the geographical name turns into a predicate (e.g., America/American), the more likely it is to be spared. 'Australia' will be retrieved more easily than 'Mount Everest' or 'The Thames'.[63]

We can observe a dissociation of impairment between common names and predicates on the one hand, and proper names on the other. In all these cases, the functioning of demonstratives and pronouns remain unaffected. From the viewpoint of Burge's theory, a case where proper names are affected while common nouns, predicates, and determiners function normally, is very hard to explain. From our viewpoint, this is a welcome piece of supporting evidence.

Recent research[64] has also shown that proper names are more likely to be affected by naturally occurring retrieval blocks than other kinds of words. This means that even in healthy speakers, proper names may be harder to recall than other kinds of words. There is also a lot of anecdotal evidence pointing to age-related loss of names in both healthy subjects and in people affected by degenerative diseases, such as Alzheimer's. However, in these all cases there are a lot of complicating factors (e.g., their frequency, age of acquisition, and the role of their bearers in a person's life)[65] and the research of this particular problem is still inconclusive.

It is quite well known that proper names are especially vulnerable to the tip-of-the-tongue phenomenon (TOT), which is different from the kind of cases described above in being not a case of

---

[62]I rely on data presented by Semenza in Goodglass and Wingfield, 1997, 115-132, and Lyons, Henley, and Kay, 2002.

[63]To put it more precisely, the more common/frequent is the adjective derived from the geographical name, the more likely is the name to be intact.

[64]For an overview see Semenza, 1997, 117-118.

[65]Some theorists hold that some names are more likely to be spared because they were acquired early, while other claim that names that are more frequent are more likely to be left intact. The jury is still out on this issue because the two explanations both rely on largely coinciding data. See Goodglass, 1993.

decay of proper names but just a temporary blockage or delay in retrieval. The tip-of-the-tongue phenomenon has been studied by Burke, MacKay, and Wade, 1991, among others. The issues connected with TOT have prompted the some reasearchers to develop an explanation in terms of an interactive activation theory of language, better known as the Node Structure Theory (NST).[66] According to NST, the activation of the lexical node for a common name (like 'baker') can benefit from multiple converging semantic connections with the semantic system. That is why common nouns are relatively less vulnerable to TOT than proper names are. The activation of a lexical node representing a proper name (like 'Baker') is thought to spread from the semantic system to the lexical node only via propositional nodes for specific individuals (that is, via connections between a name and information concerning a particular person, e.g., 'Mary Baker', 'John Baker', etc.). These propositional nodes may receive any amount of converging semantic information about the individual, but there will still be only a single, therefore vulnerable, connection in their output for the activation of the phonological form in the corresponding lexical node. That is why even if the bearer of the name is highly familiar, his or her name (e.g., Baker) is more prone to TOT than his or her occupation (e.g., baker). This explanation also applies to the so-called baker-Baker paradox, which can be observed in tasks that require learning names and occupations belonging to unfamiliar faces. It has been demonstrated repeatedly that the name 'Baker' presented as a proper name is harder to recall than the same word presented as an occupation (baker).[67] All these phenomena are compatible with the idea that the link between a proper name and its referent is a weak and arbitrary one.

None of these data present a knockdown argument against Burge's theory. If a name was indeed composed of a demonstrative and a predicate, and that predicate was a 'simple' one, that is, not further analysable, all of these phenomena could still occur even though

---

[66]See MacKey, 1987.

[67]The data in this paragraph is, again, drawn from Semenza, 1997, 117-118.

the sense in which such an unanalysable predicate would still be a predicate is dubious. Moreover, we have shown that there are some problems with the functioning of the demonstrative part of Burge's analysis. In particular, if we assume with Burge that the demonstrative in question is 'that', then the problem was to explain why a proper name *preceded* by 'that' tends to have a different referent than a proper name without it. When we look at the claim that the fact that in some languages a proper name is preceded by a definite article supports Burge's theory, we found problems with that as well. Burge's explanation seems counterintuitive to speakers of those languages (for example Greek or German). Looking at examples from English, we found again that the behaviour of names with and without a definite article and with or without a demonstrative is markedly different. This leaves the advocates of Burge's theory with the task of explaining why and when the demonstrative part of a proper name surfaces, and why proper names (at least in English) behave differently when preceded by a definite article or a demonstrative.

We then looked at data pointing to differences in the functioning of proper names and common nouns in the brain. We saw that the functioning of proper names can be impaired even where common nouns are well preserved. Burge's advocates could say that this is compatible with their theory because the kind of predicates or common nouns they had in mind was very special. They could claim that proper name predicates are susceptible to specific damage (anomia) because, being devoid of semantic content, they form a special kind of expressions. In other words, advocates of Burge's theory could even accept the NST explanation of the functioning of proper names if they modified their theory enough. We should note, however, that Burge's theory modified this extensively is almost equivalent to the ambiguity theory we have been defending. It becomes very hard to see in what sense his proper names are like predicates or common nouns once he accounts for all the facts that highlight their differences from these two kinds of terms.

6.10 Conclusion

The goal of this chapter was to investigate the indexical, demonstrative or ambiguous nature of proper names. We have examined some linguistic data showing that names are context dependent, and then turned our attention to the two basic lines of explanation of this phenomenon. The indexical theory, first advocated by Burge and later defended by Recanati, Haas-Spohn, and others, proved to be more resilient and flexible than one might have expected. We have, however, built on Kaplan's approach, and developed an explanation of ontology of proper names that enabled us to account for some less-common uses at least as well as the indexical approach. At the same time, our ontological framework has made more plausible the claim that names are ambiguous rather than indexical. Our approach also enabled us to reconcile the rigid designation claim with the fact that names (in Kripke's sense) have multiple bearers.

We used again some parts of Stalnaker's apparatus together with some insights from Grice's speech-act theory, this time to account for the resolution of names in a given context. The concept of context I used does not help us decide between the two competing theories (indexical and ambiguity approach), as it could be easily accommodated within either, but does add plausibility to an attractive way of thinking about resolution of names. It also seems that the only way in which the advocates of the indexical theory could account for some of the examples would be by indexing names, which would a move towards an ambiguity view. The data from aphasia support, in my view, the ambiguity view rather strongly. Yet even here, the indexicality view could be modified so as to accommodate them. However, I hope to have convinced the reader that a version of the indexicality view that is modified enough to account for all the data I have collected is as good as indistinguishable from the view I advocate. I hope to have shown that the ambiguity view, up till now something of a Cinderella of the proper name arena, is much more attractive than one might think at first sight, and that it deserves close attention.

Chapter 7
Final Review

At the outset we outlined a plan that included a division of this work in two basic parts, one devoted primarily to a critical analysis of a number of proposals dealing with the semantics of proper names, the other containing my own proposal. Each of these two parts was in turn subdivided into a part dealing with the descriptive, the modal, and the foundational semantics of proper names. It could seem that this rather strict division might have led to a fragmented view of the proposals we studied, as well as of my own contribution. In the end, I dare say, the opposite turned out to be the case: just as I hoped, the segmented approach helped us to emphasise the internal cohesion of the proposals we studied, and to justify a joint approach to descriptive and modal semantics in our own approach. One may thus wonder why, if I wanted to show the cohesion of each of the positions I analysed, did I choose such a segmented presentation? It was a part of the enterprise to compare various approaches to proper names, and the division into a descriptive modal, and foundational part has been helpful. It enabled us to put side to side the answers various semanticists offer regarding the same questions, which then helped us gather the desiderata for our own proposal. On the other hand, we have time and time emphasised the connection between the parts of each of the proposals we studied.

At the beginning we also expressed a view that while the analysis of modal statements containing proper names would be treated

separately, that might have been just a provisional measure. In the
course of our work, we traced a close connection between the descrip-
tive part of a proposal and its treatment of modals. It seems that
an analysis of modal statements serves to motivate, illustrate, and
clarify the descriptive part of a semantic proposal. That is also why
in our own proposal, we treated the two – descriptive semantics and
semantics of modals – together. In the background of this project
has been a question that kept on surfacing in various forms: What is
the relation between the semantics of proper names and the ontology
of the bearers of proper names? To put it differently, what is the
relation between a theory that attempts to capture the behaviour of
proper names in natural language and the predictions or claims it
makes about the identity criteria of the referents of proper names?
Is it possible for a theory to be neutral on this issue? Answers to
these questions emerged gradually. We saw that every one of the
proposals we studied did, in the end, make claims or had direct
consequences for a treatment of individuals. Even Lewis's theory,
which was designed to avoid making any claims regarding necessary
properties of individuals, ends up having direct consequences in that
area.

In the preface, I admitted that origins of the current project
date back to my reading of Kripke's *Naming and Necessity*. I had
been dissatisfied with the way Kripke connects what is presented
as descriptive semantics with substantive essentialist claims. In my
treatment of Kripke's semantics of modal statements I tried to get
to the bottom of this issue. That proved to be rather difficult. It
was relatively easy to show that the principles Kripke uses to de-
rive essentialist claims (principles such as 'it is an essential property
of an object that it is composed of the material it is actually com-
posed of'), being admittedly results of philosophical reflection, do
not originate in the semantics of the terms involved, that is in the
semantics of natural kind terms and proper names. It was less easy
to show why the kind of essentialism Kripke proposes is undesirable:
Kripke does not actually provide an argument for his position, and
in absence of an argument that could be challenged, the danger is
that we could do no better than saying that my opinion differs from

Kripke's. That, however, would be unsatisfactory. That is why I tried to point out some counterintuitive consequences of Kripke's view. While this was the second best option, it was probably the best kind of move one can make in such a similar situation.

Having rejected Kripke's essentialism as not only unsupported but also undesirable, we had to face the fact that the problems we encountered in Lewis's system were caused by a radical *lack* of any essentialist commitments. In Lewis's proposal, the absence of essentialist commitments was the root of the problems.[1] It led to the adoption of worldbound individuals related by a counterpart relation, which in turn resulted in a counterintuitive analysis of modal statements involving individuals. A further unpleasant consequence of this was the collapse of the notion of rigid designation in Lewis's framework.

What we wanted was an approach that would allow for transworld individuals and make a good sense of rigid designation, while avoiding essentialist commitments of the kind we criticised Kripke for. This is the point where the insights we gathered from our analysis of Stalnaker's work stood us in good stead. We had argued that Stalnaker's possible worlds are inherently contextdependent and that what they model is the beliefs of conversation participants. We took this basic idea one step further and applied it to the issues of reference of proper names. Within a Stalnaker-style framework, we argued, one cannot make a good sense of the notion of an absolute domain of possible worlds in which a proper name would rigidly refer – reference of rigid designators is therefore always to some degree context-dependent. One should still think of names as rigidly designating, but the kind of domain in which they refer is very different from that which Kripke had in mind. Also, we had shown that rigid designation had to be underpinned by some sort of essentialism. The essentialist commitment in question could be purely formal, as in Kaplan's case, or rather substantive, as in Kripke's case. Neither of these options seemed very attractive because neither seems to

---

[1]Essentialism could be added to Lewis's possible world framework, it is general enough. But even so it is unclear whether that would lead to an adoption of transworld individuals.

capture adequately speaker's intuitions as what they refer to. However, once we started working with context-dependent domains of epistemic possible worlds, we could argue that the individuals have properties that are necessary of them in a given context. This is a sort of essentialism that does, indeed, seem to capture the way names are used in natural language. The sort of weak essentialism with which we underpin rigid designation seems to avoid both the counterintuitive consequences of Kripke's essentialism and the purely formal character of Kaplan's haecceitism.

In this way we progressed from an analysis of preconditions of the direct reference and rigid designation picture to an analysis of the various notions of possible worlds and the way they influence the descriptive semantics of proper names, and, finally, put the pieces together and formed our own proposal.

In the course of our work, we tried to avoid the problems connected with causal chains. In both Kripke's and Putnam's proposal, causal chains seem to play the role of truth-warrants. As long as we could account for the truth-conditions of sentences containing proper names, we could therefore avoid the issue of causal chains. Instead of analysing various proposals dealing with causal chains, we turned our attention to Stalnaker's information theoretic account of intentionality, which is designed to explain the way in which the content of speaker's statements is connected to the external world. We found Stalnaker's approach plausible, though it may profit from further elaboration.

Finally, we addressed the issues of foundational semantics of proper names, in particular the problems connected with the resolution of names in context. With the exception of Kaplan's work, none of the proposals we had analysed up to then dealt with this problem. That is why the last chapter seems to some degree disconnected from the rest of the work. We felt, however, that the issue of multiple bearers of proper names had to be addressed, if only because it had a bearing on the rigid designation of proper names and the ontology of names. The last chapter also connects with the rest of the work by asking and answering the questions like 'What is a proper name?' and 'What is special about proper names?'. Tak-

ing Kaplan's proposal as our starting point, we proposed a view of the ontology of proper names and reconciled the rigid designation claim with the problems surrounding multiple bearers of the same name. Using data from aphasia research, we argued against Burge and others that names are indeed a category of their own, and that they cannot be analysed in terms of other kinds of expressions.

There are issues that could have been addressed in the present work and we chose not to deal with them, e.g., the topic of causal chains. I hope, though, that the reasons behind the choice of topics we did address have become as apparent to the reader as they were to me when I made the choice. I hope, too, that this work presents not only an interesting analysis of connections between various issues present in a theory of reference of proper names, but also some refreshing and thoughtprovoking new perspectives.

Bibliography

Adams, Robert Merrihew, 1979, Theories of Actuality. In: Michael J. Loux (ed.), *The Possible and the Actual*, chap. 10, 190–210, Ithaca: Cornell University Press.

Alward, Peter, 2004, Varieties of Believed-World Semantics, http://people.uleth.ca/ peter.alward/papers.

Bach, Kent, 2002a, Seemingly Semantic Intuitions. In: Keim J. Campbell, M. O'Rourke, and D. Shier (eds.), *Meaning and Truth*, 21–33, New York: Seven Bridges Press.

———, 2002b, Semantic, Pragmatic. In: Keim J. Campbell, M. O'Rourke, and D. Shier (eds.), *Meaning and Truth*, 284–292, New York: Seven Bridges Press.

Bednar, James A. and Miikkulainen, Risto, 2000, Self-Organisation of Innate Face Preferences: Could Genetics Be Expressed Through Learning? In: James Bednar (ed.), *Proceeding of the Seventeenth National Conference on Artificial Intelligence*, 117–122, Cambridge, Massachusetts: MIT Press.

Benton, Arthur L. and Joynt, Robert J., 1960, Early Descriptions of Aphasia. *Archives of Neurology*, **3**: 109–126.

Bilgrami, Akeel, 1992, *Belief and Meaning*. Oxford: Basil Blackwell.

———, 1998, Precis of Belief and Meaning. *Philosophy and Phenomenological Research*, **LVIII**(3): 595–605.

Block, Ned, 1986, An Advertisement of a Semantics for Psychology. *Midwest Studies in Philosophy*, **10**: 615–678.

Bower, Bruce, 2002, Baby Facial: Infants Monkey with Face Recognition. *Science News*, **161**(20): 305–307.

Burge, Tyler, 1973, Reference and Proper Names. *The Journal of Philosophy*, **70**: 425–438.

———, 1979, Individualism and the Mental. *Midwest Studies in Philosophy*, **4**: 73–121.

Burke, Deborah M., MacKay, Don G., and Wade, Julia S., 1991, On the Tip of the Tongue: What Causes Word Finding Difficulties in Young and Older Adults. *Journal of Memory and Language*, **30**: 542–579.

Caplan, Ben, 2004, Putting Things in Context. Philosophical Review, forthcoming.

Capone, Alessandro, 2003, Common Ground (Again!). Online Papers in Philosophy.

Carlson, Greg, 1977, *Reference to Kinds in English*. Ph.D. thesis, Graduate School of the University of Massachusetts, Amherst, MA.

Carnap, Rudolf, 1947, *Meaning and Necessity*. Chicago, Illinois: University of Chicago Press.

Carver, L. J. and Dawson, G., 2002, Anatomy and Neurology of Autism: Development and Neural Bases of Face Recognition in Autism. *Molecular Psychiatry*, **7**: 18–20.

Cherniak, Christopher, 1986, *Minimal Rationality*. MIT Press.

Chisholm, Roderick M., 1979, Identity Through Possible Worlds. In: Michael Loux (ed.), *The Possible and the Actual*, 80–88, Ithaca: Cornell University Press.

Cresswell, Max J., 1985, *Structured Meanings*. Cambridge, Massachusetts: MIT Press.

Davidson, Donald, 1968, On Saying That. *Synthese*, **19**.

Donnellan, Keith, 1966, Reference and Definite Descriptions. *The Philosophical Review*, **75**(281–304).

———, 1970, Proper Names and Identifying Desriptions. *Synthese*, **21**: 335–58.

Dummett, Michael, 1973a, *Frege: Philosophy of Language*. London: Duckworth.

———, 1973b, *Frege: Philosophy of Language*. London: Duckworth.

———, 1973c, *The Seas of Language.* Clarendon Press: Oxford.

Ebbs, Gary, 1998, Bilgrami's Theory of Belief and Meaning. *Philosophy and Phenomenological Research*, **LVIII**(3): 613–620.

Evans, Gareth, 1973, The Causal Theory of Names. *Aristotelian Society Proceedings, suplementary volume*, **47**: 187–208.

———, 1982, *Varieties of Reference.* Oxford: Clarendon Press.

Field, Hartry, 1977, Logic, Meaning, and Conceptual Role. *Journal of Philosophy*, **74**: 379–409.

Fodor, Jerold, 1981, Propositional Attitudes. In: Jerold Fodor (ed.), *RePresentations*, 177–203, Cambridge, Massachusetts: MIT Press.

———, 1987, *Psychosemantics.* Cambridge, MA: MIT Press.

Fodor, Jerry and LePore, Ernesto, 1992, *Holism: A Shopper's Guide.* Oxford: Basil Blackwell.

Forbes, Graeme, 1990, The Indispensability of Sinn. *The Philosophical Review*, **99**: 535–564.

———, 1996, Substitutivity and the Coherence of Quantifying In. *The Philosophical Review*, **105**: 337–372.

Frege, Gottlob, 1893, re-print in English 1952, On Sense and Reference. In: Peter Geach and Max Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege*, Oxford: Basil Blackwell, first published in Zeitschrift fuer Philosophie und Philosophische Kritik, pg. 25 - 50, as Ueber Sinn und Bedeutung.

French, P. A., Uehling, T. E., and Wettstein, H. K. (eds.), 1986, *Studies in Essentialism*, Midwest Studies in Philosophy, XI, Minneapolis: University of Minnesota Press.

Geach, Peter, 1968, *Reference and Generality.* Ithaca: Cornell University Press.

Gibbs, Raymond and Moise, Jessica, 1997, Pragmatics in Understanding What Is Said. *Cognition*, **62**: 51–74.

Goodglass, Harold, 1993, *Understanding Aphasia.* San Diego: Academic Press.

Goodglass, Harold and Wingfield, Arthur (eds.), 1997, *Anomia – Neuroanatomical and Cognitive Correlates.* San Diego: Academic Press.

Grice, Paul, 1975, Logic and Conversation. In: P. Cole and J. Mor-

gan (eds.), *Syntax and Semantics*, vol. 3, 41–58, New York: Academic Press.

Guttenplan, Samuel (ed.), 1995, *A Companion to the Philosophy of Mind*. Oxford: Basil Blackwell.

Haas-Spohn, Ulrike, 1994, *Hidden Indexicality and Subjective Meaning*. Ph.D. thesis.

Harman, Gilbert, 1973, *Thought*. Princeton: Princeton University Press.

———, 1986, *Change in View: Principles of Reasoning*. Bradford Books, Cambridge, Mass.: MIT Press.

Harper, William L., Stalnaker, Robert, and Pearce, Glenn (eds.), 1981, *Ifs*. Dordrecht: Reidel.

Heim, Irene, 1988, *The Semantics of Definite and Indefinite Noun Phrases*. New York: Garland.

Hume, David, 1748, *Enquiry Concerning Human Understanding*. Indianapolis: Library of the Liberal Arts, edited by charles w. hendel, 1955 ed.

Kamp, Hans, 1981, A Theory of Truth and Semantic Represenation. In: Jeroen Groenendijk, Theo Janssen, and Martin Stokhof (eds.), *Formal Methods in the Study of Language*, 277– 322, Amsterdam: Mathematisch Centrum.

Kamp, Hans and Reyle, Uwe, 1993, *From Discourse to Logic: An Introduction to Modeltheoretic Semantics Of Natural Language, Formal Logic and DRT*. Dordrecht: Kluwer.

Kant, Immanuel, 1787, 1963 edition, *Critique of Pure Reason*. London: Macmillan and Co. Ltd., 1963 ed., translated by Norman Kemp Smith.

Kaplan, David, 1973, Bob and Carol and Ted and Alice. In: K.J.J. Hintikka, J.M.E. Moravczik, and P. Suppes (eds.), *Approaches to Natural Language*, 490–519, Dordrecht: Reidel.

———, 1975, How to Rusell a Frege–Church. *The Journal of Philosophy*, **72**(716–729).

———, 1978, On the Logic of Demonstratives. *Journal of Philosophical Logic*, **8**: 81–98.

———, 1979, Dthat. In: P.A. French, T.E. Uehling, and H.K. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of*

Language, 383–401, Minneapolis: University of Minnesota Press.

———, 1979a, On the Logic of Demonstratives. In: P.A. French, T.E. Uehling, and H.K. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, 401–414, University of Minnesota Press.

———, 1979c, Transworld Heir Lines. In: Michael Loux (ed.), *The Possible and the Actual*, 88–109, Ithaca: Cornell University Press.

———, 1989a, Demonstratives. In: J. Almog, J. Perry, and H.Wettstein (eds.), *Themes from Kaplan*, 481–563, Oxford: Oxford Univerity Press.

———, 1989b, Afterthoughts. In: J. Almog, J. Perry, and H.Wettstein (eds.), *Themes from Kaplan*, 565–614, Oxford: Oxford Univerity Press.

———, 1990, Words. *The Aristotelian Society Proceedings, supplementary volume*, **64**: 93–119.

King, Jeffrey C., 1996, Structured Propositions and Sentence Structure. *Journal of Philosophical Logic*, **29**(4): 516 – 535.

Kripke, Saul, 1963, Semantical Considerations on Modal Logic. *Acta Philosophica Phennica*, **16**: 83–94.

———, 1971, Identity and Necessity. In: Munitz and Milton (eds.), *Identity and Individuation*, New York: New York University Press.

———, 1972a, Addenda to Naming and Necessity. In: Donald Davidson and Gilbert Harman (eds.), *S*, 763–769, Cambridge, Mass.: Semantics of Natural Language.

———, 1977a, Identity and Necessity. In: Stephen P. Schwartz (ed.), *Naming, Necessity, and Natural Kinds*, 66–101, London: Cornell University Press.

———, 1977b, Speaker's Reference and Semantic Reference. *Midwest Studies in Philosophy*, **2**: 255–276.

———, 1979, A puzzle about belief. In: A. Margalit (ed.), *Meaning and Use*, Dordrecht: Reidel.

———, 1980, *Naming and Necessity*. Cambridge. Mass.: Harvard University Press.

Lepore, Ernest and Cappelen, Herman, 1997, Varieties of Quotation. *Mind*, **106**: 429–450.

Lepore, Ernest and Fodor, Jerry, 1992, *Holism: A Shopper's Guide.*

Oxford: Basil Blackwell.

Lerner, Jean-Yves and Zimmerman, Ede, 1984, Bedeutung und Inhalt Von Eigennahmen. SBF 99, papier Nr. 94, Konstanz.

Lewis, David, 1972, General Semantics. In: Donald Davidson and Gilbert Harman (eds.), *Semantics of Natural Language*, Dordrecht: Reidel.

———, 1973, *Counterfactuals*. Cambridge, Mass.: Cambridge University Press.

———, 1979a, Possible Worlds. In: Michael Loux (ed.), *The Possible and the Actual*, 182–190, Ithaca: Cornell University Press.

———, 1979b, Counterpart Theory and Quantified Modal Logic. In: Michael Loux (ed.), *The Possible and the Actual*, 110–129, Ithaca: Cornell University Press.

———, 1986, *On the Plurality of Worlds*. Oxford: Basil Blackwell.

———, 1995, David Lewis: Reduction of Mind. In: Samuel Guttenplan (ed.), *A Companion to the Philosophy of Mind*, 412–431, Oxford: Basil Blackwell.

Linsky, Leonard, 1977, *Names and Descriptions*. Chicago: University of Chicago Press.

Loux, Michael L. (ed.), 1979, *The Possible and the Actual (Readings in the Metaphysics of Modality)*. Ithaca: Cornell University Press.

Lowe, E. Jonathan, 1997, Objects and Criteria of Identity. In: Bob Hale and Crispin Wright (eds.), *A Companion to the Philosophy of Language*, 613–634, Oxford: Basil Blackwell.

Lycan, William, 1979, The Trouble with Possible Worlds. In: Michael Loux (ed.), *The Possible and the Actual*, 274–317, Ithaca: Cornell University Pres.

Lyons, Frances, Henley, Richard J., and Kay, Janice, 2002, Anomia for Common Names and Geographical Names with Preserved Retrieval of Names of People: A Semantic Memory Disorder. *Cortex*, **38**: 23–35.

MacKey, Don G., 1987, *The Organisation Of Perception and Action: A Theory for Language and Other Congnitive Skills*. New York: Springer Verlag.

Marcus, Ruth Barcan, 1961, Modalities and Intensional Languages. *Synthese*, **13**: 303–322.

McGinn, Colin, 1976, On the Necessity of Origin. *Journal of Philosophy*, **73**: 127–35.

Mill, John Stuart, 1872, reprint 1949, *A System of Logic*. London: Longmans, Green and Co., definiftive 8th ed.

Moore, Joseph G., 1999, Misdisquotation and Substitutivity: When Not to Infer Belief from Assent. *Mind*, **108**: 335–365.

Moore, Robert, 1988, Is It Rational to Be Logical? In: M. Y. Vardi (ed.), *Proceedings of the Second Conference on Theoretical Aspects of Reasoning About Knowledge*, Los Altos, California: Morgan Kaufman.

Morris, William (ed.), 1981, *The American Heritage Dictionary of the English Language*. Boston: Houghton Mifflin Company, 10th ed.

Neale, Stephen, 1990, *Descriptions*. Cambridge, Massachusetts: The MIT Press.

Ostertag, Gary (ed.), 1998, *Definite Descriptions, A Reader*. A Bradford Book, Cambridge, Massachusetts: MIT Press.

Pelczar, Michael and Rainsbury, Joseph, 1998, The Indexical Character of Names. *Synthese*, **114 (2)**: 293–317.

Perry, John, 1980, A Problem About Continued Belief. *Pacific Philosophical Quarterly*, **5**: 533–542.

———, 1997, Indexicals and Demonstratives. In: Bob Hale and Crispin Wright (eds.), *A Companion to the Philosophy of Language*, 586–613, Oxford: Basil Blackwell.

Poesio, Massimo and Traum, David, 1997, Representing Conversation Acts in a Unified Semantic/Pragmatic Framework. draft.

Putnam, Hilary, 1975a, The Meaning of 'Meaning'. In: *Philosophical Papers*, vol. 2, Mind, Language, and Reality, 215–71, Cambridge: Cambridge University Press.

———, 1975b, *Philosophical Papers: Mind, Language, and Reality*, vol. 2. Cambridge: Cambridge University Press.

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

Quine, Willard Van Orman, 1960, *Word and Object*. Cambridge, MA: MIT Press.

———, 1961a, On What There Is. In: *From A Logical Point of*

View, chap. I, 1–20, New York: Harper and Row.

———, 1961b, Two Dogmas of Empiricism. In: *From a Logical Point of View*, chap. II, 20–47, New York: Harper and Row.

Recanati, Francois, 1989, The Pragatics of What is Said. *Mind and Language*, **4**: 294–328.

———, 1993, *Direct Reference: From Language to Thought*. Oxford: Basil Blackwell.

Richards, Tom, 1975, Worlds of David Lewis. *Australasian Journal of Philosophy*, **53**: 105–118.

Roberts, Craige, 1996, Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. In: Jae Hak Yoon and Andreas Kathol (eds.), *Papers in Semantics*, vol. 49, http://www.semanticsarchive.net/.

Rosenberg, Jay F., 1994, *Beyond Fromalism (Naming and necessity for Human Beings)*. Philadelphia: Temple University Press.

Russell, Sir Bertrand, 1905, On Denoting. *Mind*, **14**: 479–493, reprinted in Definite Descriptions:A reader, Gary Ostertag (ed.), The MIT Press, Cambridge Massachusetts, 1998.

———, 1911, Knowledge by Acquitance and Knowledge by Description. In: Sir Bertrand Russell (ed.), *Mysticism and Logic and Other Essays*, 209–232, London: Longmas, Green and Co.

———, 1956, The Philosophy of Logical Atomism. In: Robert Charles Marsh (ed.), *Bertrand Russell: Logic and Knowledge, Essays 1901–1950*, 175–283, London: George Allen and Unwin.

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

———, 1986, *Frege's Puzzle*. cambridge, MA: MIT Press.

Searle, John, 1967, Proper Names. In: Sir Peter Strawson (ed.), *Philosophical Logic*, Oxford: Oxford Univerity Press.

Segal, Gabriel, 2001, Two Theories of Names. *Mind and Language*, **5**: 547–563.

Segal, Gabriel and Larson, Richard, 1995, *Knowledge of Meaning: An Introduction to Semantic Theory*. Cambridge, MA: MIT Press.

Semenza, Carlo, 1997, Proper-Name-Specific Aphasias. In: H. Goodglass and A. Wingfield (eds.), *Anomia, Neuroanatomical and Cog-*

nitive Correlates, San Diego: Academic Press.

Smith, Quentin, 1995, Marcus, Kripke, and the Origin of the New Theory of Reference. *Synthese*, **104**(2): 179–189.

Sosa, David, 1996, The Import of the Puzzle About Belief. *The Philosophical review*, **105**: 373–402.

Stalnaker, Robert, 1978, Assertion. *Syntax and Semantics*, **9**: 315–32.

———, 1979, Possible Worlds. In: Michael Loux (ed.), *The Possible and the Actual*, 225–234, Ithaca: Cornell University Press.

———, 1981, Indexical Belief. *Synthese*, **49**: 129–151.

———, 1984, *Inquiry*. Cambridge, Massachusetts: MIT Press.

———, 1987, Semantics for Belief. *Philosophical Topics*, **15**: 177–190.

———, 1989, On What's in the Head. In: *Philosophical Perspectives, 3: Philosophy of Mind and Action Theory*, Oxford: Oxford University Press.

———, 1990, Narrow Content. In: C.A. Anderson and J. Owens (eds.), *Propositional Attitudes: The Role of Content in Logic*, 131–145, Stanford: CSLI.

———, 1991, How to Do Semantics for the Language of Thought. In: B. Loewer and G. Rey (eds.), *Meaning and Mind – Fodor and His Critics*, 229–319, Oxford: Basil Blackwell.

———, 1993, Twin Earth Revisited. *Proceedings of the Aristotelian Society*, **XCIII**: 297–311.

———, 1997, Reference and Necessity. In: Bob Hale and Crispin Wright (eds.), *A Companion to the Philosophy of Language*, 534–555, Oxford: Basil Blackwell.

———, 1998, On the Representation of Context. *Journal of Logic, Language, and Information*, **7**: 3–19.

———, 1999a, Belief Attribution and Context. In: Robert Stalnaker (ed.), *Context and Content*, 150 – 167, Oxford: Oxford University Press.

———, 1999b, *Context and Content*. Oxford: Oxford University Press.

———, 1999c, The Problem of Logical Omniscience, I. In: Robert Stalnaker (ed.), *Context and Content*, 241 – 255, Oxford: Oxford

University Press.

———, 1999d, The Problem of Logical Omniscience, II. In: Robert Stalnaker (ed.), *Context and Content*, 255 – 274, Oxford: Oxford University Press.

Stanley, Jason, 1997, Names and rigid designation. In: Bob Hale and Crispin Wright (eds.), *A Companion to the Philosophy of Language*, Oxford: Basil Blackwell.

Strawson, Sir Peter F. (ed.), 1967, *Philosophical Logic*. Oxford: Oxford University Press.

Strawson, Sir Peter Frederick, 1950, re-printed in 1998, On Referring. In: Gary Ostertag (ed.), *Definite Descriptions, A Reader*, 135–161, Cambridge, Massachusetts: MIT Press.

———, 1959, *Individuals*. London: Methuen.

Valentin, Dominique and Abdi, Hervao, 2001, Face Recognition by Myopic Baby Neural Networks. *Infant and Child Development*, **10**(1): 19–20.

Wiggins, David, 1980, *Sameness and Substance*. Cambridge, Massachusetts: Harvard University Press.

Wright, Crispin, 1993, *Realism, Meaning and Truth*. Oxford: Basil Blackwell.

Name Index

Subject Index

Nederlandstalige Samenvatting

Dit werk vindt zijn oorsprong in het lezen van Kripkes *Naming and Necessity* in 1993. Deze lectuur had in mij een gevoel van onvrede nagelaten dat lang genoeg bleef hangen om als inspiratie te dienen voor mijn MA-scriptie (over *Internalism and Externalism in the Theories of Reference*) en nu ook voor dit proefschrift. In de tussenliggende periode maakte ik kennis met andere invloedrijke benaderingen van de referentie van eigennamen, maar mijn onvrede met het essentialisme en mijn wantrouwen jegens directe verwijzing bleef bestaan.

In dit proefschrift met de titel *A Users Guide to Proper Names* heb ik een poging ondernomen om te analyseren wat er – volgens mij – problematisch is aan de geaccepteerde theorieën over de referentie van eigennamen. Ik heb geprobeerd een aantal aan elkaar gerelateerde doctrines over het gedrag van eigennamen uit elkaar te halen om zo tot een beter begrip te komen van de wijze waarop de verschillende delen van de referentietheorieën die ik wil behandelen met elkaar verbonden zijn. Dit heeft het mogelijk gemaakt om mijn eigen voorstel over de semantiek en pragmatiek van eigennamen te ontwikkelen.

De *Guide* is opgebouwd rond een specifieke benadering van de taken van een semantische theorie. Volgens deze aanpak, voorgesteld door Stalnaker, moet een semantische theorie over eigennamen rekenschap geven van de descriptieve semantiek van eigennamen, van hun principiële semantiek en van de semantiek van de modale uitspraken

waarin ze voorkomen. Descriptieve semantiek gaat over de bijdrage van een eigennaam aan de waarheidswaarde van zinnen waarin de naam voorkomt. Op basis van zo'n analyse krijgt de eigennaam een semantische waarde die verondersteld wordt ons een interpretatie van die naam te geven. Een centraal deel van deze taak bestaat eruit duidelijk te maken wat voor soort ding de semantische waarde van een eigennaam is.

Bij het interpreteren van zinnen waarin eigennamen voorkomen kan gebruik worden gemaakt van het begrip van mogelijke werelden. Dit gebeurt vaak en natuurlijk in het bijzonder waar het gaat om modale beweringen. Er zijn verschillende benaderingen van modaliteit die verschillende groepen vooronderstellingen met zich meebrengen. Een analyse van de systematische eigenschappen en vooronderstellingen van verschillende manieren waarop een mogelijke-wereldensemantiek kan worden opgezet, is een deeltaak van de semantiek van modale uitspraken. Een andere taak van de semantiek van modale uitspraken betreft het onderzoek naar de oorsprong van de beperkingen op mogelijke werelden die gebruikt worden in de analyse van deze uitspraken, dat wil zeggen of, en tot op welke hoogte, ze afgeleid kunnen worden uit het descriptieve of principiële gedeelte van de semantische taakstelling.

Beide gedeelten van een semantische theorie maken potentieel voorspellingen over de principiële semantiek van eigennamen, die gaat over het gedrag van de spreker en over communicatie. Principiële semantiek beoogt een antwoord te geven op de volgende vraag: wat ligt er ten grondslag aan het feit dat de taal die door een individu of door een gemeenschap wordt gesproken juist die descriptieve semantiek heeft welke zij heeft? In dit onderzoek bekijkt men de spreker, haar bedoelingen en communicatiedoelen en probeert de strategien te ontdekken die zij gebruikt om haar (talige) boodschap over te brengen.

De *Users Guide to Proper Names* bestaat uit twee delen. Het eerste gedeelte, bestaande uit de hoofdstukken 1 tot 4, gaat over de reconstructie en analyse van een aantal invloedrijke benaderingen van de semantiek van eigennamen. Het tweede deel van het proefschrift, hoofdstuk 5 tot 7 bevat mijn eigen voorstel met betrekking

tot de semantiek van eigennamen en een conclusie.

Het eerste gedeelte van het proefschrift begint met een inleidend hoofdstuk, *Outlining the Field and Introducing Some of the Players*. Dit hoofdstuk beoogt het nut te laten te laten zien van de organisatie van het proefschrift rond het onderscheid tussen descriptieve, modale en principiële semantiek en een precieze karakterisering te geven van deze begrippen. In de tweede helft van hetzelfde hoofdstuk, *Setting the Scene*, geef ik een voorlopige schets van sommige begrippen die nodig zijn om de referentietheoriën te beschrijven die later onderzocht worden. Dat zijn voornamelijk de theoriën van Kripke, Kaplan, Lewis en Stalnaker.

In het tweede hoofdstuk, *Descriptive Semantics*, staat de descriptieve semantiek centraal zoals voorgesteld door Kripke en Kaplan voor eigennamen. In het eerste deel van het hoofdstuk behandel ik verschillende argumenten voor starre verwijzing. Nadere beschouwing leert dat in feite geen van de drie argumenten in kwestie – het modale, het kentheoretische en het semantische argument – een argument voor starre verwijzing is. Alle drie richten ze zich uitsluitend tegen bepaalde vormen van descriptivisme. Bovendien is het relatief eenvoudig, in het geval van het modale en dat van het kentheoretische argument, om formuleringen van descriptivisme te vinden die immuun zijn voor de wijze van redeneren in deze argumentaties. Het semantische argument lijkt het sterkste van de drie omdat het berust op een directe botsing van intuïties met betrekking tot de identiteit van individuele objecten.

Een analyse van een direct argument voor starre verwijzing vestigt tevens de aandacht op de verbinding tussen starre verwijzing en bepaalde vooronderstellingen over de identiteit van individuele objecten. Een nader onderzoek van de vragen met betrekking tot de identiteit van individuele objecten in verschillende mogelijke werelden blijkt op dit punt geboden. Het wordt ook duidelijk dat starre verwijzing niet volstaat om Kripkes opvatting van namen als nondescriptieve entiteiten die refereren zonder bemiddeling van enige conceptuele inhoud volledig te bepalen.

In het tweede deel van hoofdstuk 2 introduceer ik de grondbegrippen van Kaplans benadering van de descriptieve semantiek van

eigennamen. Om aan te komen bij Kripkes bekende opvatting van eigennamen blijkt het noodzakelijk om directe referentie voor namen en tenminste één versie van haecceïtisme te vooronderstellen. Hoewel – zoals ik laat zien – Kaplans opzet niet goed werkt voor eigennamen, geeft deze opzet wel een aantal conceptuele hulpmiddelen die nuttig zijn voor de latere hoofdstukken.

In het derde hoofdstuk, *Modal Statements, Individuals, and Essences*, introduceer ik drie verschillende benaderingen om een mogelijke-wereldensemantiek op te zetten, die van Lewis, die van Kripke en die van Stalnaker. In alle drie gevallen worden dezelfde vragen gesteld: Wat is de motivering en het beoogde terrein van toepassing van deze opzet? Wat zijn de ontologische gevolgen van de benadering? Welke vorm van essentialisme (als daarvan sprake is) is het gevolg? Hoe gaat de benadering om met de notie van een individueel object?

Het voornaamste probleem bij Lewis blijkt de theorie van counterparts te zijn die zoals ik laat zien een wezenlijk bestanddeel is van zijn aanpak. Het begrip van een individueel object dat erdoor wordt geïmpliceerd, komt niet overeen met enige intuïtieve lezing van contrafactische implicaties waarin verwezen wordt naar individuele objecten. Bij Kripkes theorie gaat het vooral om het vaststellen van de zwakste vorm van essentialisme waarbij het voorstel inderdaad doet wat het beoogt te doen. Hierna analyseer ik het essentialisme dat Kripke feitelijk voorstelt en de motivering en de vooronderstellingen waarop het berust. Ik concludeer dat de motivering niet voorkomt uit een analyse van natuurlijke taal en dat zij een bepaalde vorm van wetenschappelijk realisme vooronderstelt. De rest van het hoofdstuk is gewijd aan een reconstructie en analyse van Stalnakers theorie van mogelijke werelden, een theorie die aanmerkelijk voorzichtiger blijkt te zijn met betrekking tot metafysische vooronderstellingen en geschikter voor een analyse van natuurlijke taal. Net zoals bij de vorige twee voorstellen probeer ik ook hier het concept van de actuele wereld dat verondersteld wordt te reconstrueren.

In het vierde hoofdstuk, *Foundational Semantics*, onderzoek ik het concept van propositie dat geïmpliceerd wordt door de benaderingen van Lewis, Kripke en Stalnaker. Ik bekijk vooral proposities die worden uitgedrukt door zinnen waarin eigennamen voorkomen en

analyseer de manier waarop ieder van de drie concepties kwetsbaar
is voor het probleem van logische alwetendheid. In het geval van
Lewis blijkt dat het begrip propositie zo zwak is dat het nogal on-
interessant wordt. Mijn analyse van Kripkes begrip van propositie
gaat niet alleen over de systematische aspecten maar ook over Krip-
kes Pierre-probleem en een aantal pogingen om dit probleem op te
lossen. Ik bespreek een aantal benaderingen en vergelijk hun ver-
diensten. Stalnaker neemt het probleem van logische alwetendheid
zeer serieus, maar toch is zijn aanpak minder kwetsbaar voor het
probleem van logische alwetendheid dan de andere benaderingen die
ik bespreek. Ik laat zien dat een aantal pragmatische eigenschap-
pen van Stalnakers benadering (het epistemische karakter van zijn
theorie van mogelijke werelden, de Griceaanse principes die in het
begrip van assertie zijn ingebouwd) helpen bij het bieden van een
tegenwicht tot, en het afzwakken van, de omvang van het probleem
van logische alwetendheid.

In het vijfde hoofdstuk, *User-friendly Descriptive Semantics*, dat
ook het eerste hoofdstuk is van het tweede deel, introduceer en
motiveer ik mijn eigen voorstel voor de descriptieve semantiek van
eigennamen en de modale uitingen waarin eigennamen voorkomen.
Mijn voornaamste bedoeling is om het begrip van eigennamen als
starre verwijzers te handhaven onder vermijding van overbodige meta-
fysische aannames. Door te werken met epistemische mogelijke we-
relden waarvan het domein in alle gevallen medebepaald wordt door
een specifieke context kan ik een zeer intuïtief en plausibel begrip
van individueel object ontwikkelen dat als starre verwijzing voor een
eigennaam kan dienen. De interpretatie van modale uitingen wordt
dan niet gestuurd door essentialistische beperkingen (in de pretheo-
retische betekenis van deze term), maar door contextafhankelijke
beperkingen, wat een natuurlijker optie lijkt.

In het zesde hoofdstuk, *Foundational Semantics: Names, Index-
icality, and Ambiguity*, ontwikkel ik een perspectief op de principiële
semantiek die geinspireerd is door de pragmatische waarneming dat
het normale spraakgebruik toelaat om te zeggen dat een naam, b.v.
"John Smith" naar verschillende individuen kan verwijzen. Ik geef
veel aandacht aan de ontologie van eigennamen en de vraag "Wat is

een naam?" leidt tot antwoorden die vervolgens gebruikt worden in een discussie over het probleem of namen beschouwd moeten worden als indexicaal of als ambigu. Ik kies voor ambiguïteit en stel een manier voor om Stalnakers theorie van mogelijke werelden te gebruiken om intuïtief plausibele verklaringen af te leiden voor een aantal moeilijke gevallen.

In het slothoofdstuk, *Conclusion*, benadruk ik dat mijn aanpak in de hele *Guide* berust op een zorgvuldige toetsing van de uiteenlopende vooronderstellingen die aan de verschillende theorieën van mogelijke werelden ten grondslag liggen en van de uiteenlopende concepties van individueel object die gebruikt worden in de beschrijvingen van het gedrag van eigennamen. In mijn eigen benadering vermijd ik alle metafysische aanames en probeer communicatie te beschrijven in begrippen die toegankelijk zijn voor spreker en toehoorder.