

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

A Thesis Submitted for the Degree of PhD at the University of Warwick

<http://go.warwick.ac.uk/wrap/59510>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

The Dynamics and Communication of Concepts

by

Simon James Prosser

A thesis submitted in partial fulfilment of the requirements for the degree
of Doctor of Philosophy in Philosophy

University of Warwick, Department of Philosophy
November 2001

Contents

Acknowledgements	iii
Declaration	iv
Abstract	v
Introduction	1
PART 1: PRELIMINARIES	
1. The Transparency Principle	10
1.1 A criterion for sameness	12
1.2 The Transparency Principle	18
PART 2: CONCEPTS	
2. Concepts I: What Concepts Cannot Be	27
2.1 Ways of Thinking and Ways of Referring	27
2.2 Ways of Referring and the individuation of concepts	34
2.3 Varieties of Inferential Role Semantics	37
2.4 Fodor's arguments for Conceptual Atomism	43
2.5 The dynamic argument for Conceptual Atomism	48
2.6 Changing one's mind	57
2.7 Reason and perception	61
2.8 Summary	72
3. Concepts II: An Atomistic Theory	74
3.1 A convergence of theories	75
3.2 Concepts as referential episodes	86
3.3 Analyticity and holism	93
3.4 Psychology and causation	99
Appendix A: Reference, Causation and Information	101
Appendix B: Spatiotemporal Location and Reference to Physical Objects	112
4. Concepts III: Communication	116
4.1 Communication and shared concepts	116
4.2 Interpretive versus transparent communication	121
4.3 Word-using practices	126
4.4 Demonstratives and the Paderewski paradox	138

PART 3: INDEXICALS

5. Changing Indexicals	143
5.1 Evans's argument	144
5.2 Transparency	147
5.3 Further considerations	153
5.4 Diachronic anaphors	156
6. The Role of Egocentricity in Indexical Thought	158
6.1 Indexical thought and action	158
6.2 The structure of indexical thought	162
6.3 Egocentric predicates	167
6.4 The role of egocentric predicates in communication	169
6.5 Indexicals and complex demonstratives	171
6.6 Immunity to error through misidentification	174
6.7 I am here now	186
6.8 Rip van Winkle	191
Appendix C: An Argument for the Tenseless Theory of Time	194
Bibliography	201
Diagrams	
1. Square, regular diamond	65
2. Communication using demonstratives	138

Acknowledgements

The studies for this thesis were conducted with the help of a three year full-time studentship from the Arts and Humanities Research Board (AHRB). I would like to thank the AHRB for this support, without which my studies might well have been impossible. I am also grateful to the Department of Philosophy at the University of Warwick for helping me in numerous ways over the years and to the Entente Cordiale Trust for financial support during a year spent CREA (CNRS/École Polytechnique) in Paris which, although not strictly part of my doctoral studies, no doubt had a beneficial influence on them. I would also like to thank CREA and its staff (particularly those who have now moved to the Institut Jean Nicod) for their friendly support during that year.

I am enormously grateful to Michael Luntley, whose supervision of this thesis has been absolutely first rate from beginning to end. Tales from other students at various institutions have often made me realise how fortunate I have been to have such a dedicated and professional supervisor. Discussing my ideas with Michael has always been an enormous pleasure and I have learned a very large proportion of what I know from him.

A number of people have read drafts of sections of this thesis and their comments have been extremely helpful. In particular I would like to thank José Luis Bermúdez, Jérôme Dokic, Naomi Eilan, Christoph Hoerl, François Recanati and Johannes Roessler. I would also like to thank Stacie Friend, Lynn Guyver and Aleka Lianeri. Finally, I would like to thank my parents, Tony and Joan Prosser, for their endless unconditional support over a number of difficult years.

Declaration

The work contained in this thesis is entirely the author's own work and has not been submitted for any other degree or at any other university. None of the material has been published elsewhere apart from some sections of Appendix C which are taken from the author's paper 'A New Problem for the A-Theory of Time', *The Philosophical Quarterly*, 50 (2000), pp. 494-498. This material was written in the course of the author's doctoral studies and is therefore included within the work to be examined for the degree.

Simon Prosser

November 2001

Abstract

The central claim of this thesis is that concepts, the components from which cognitively significant truth evaluable content (thought) is composed, are unstructured entities an account of whose individuation makes no essential reference to other concepts in the possession of the thinking subject or to any particular means by which the reference of the concept is identified by the thinking subject. This position is called Conceptual Atomism and contrasts with Inferential Role Semantics, according to which concepts are individuated by their inferential roles or their conditions of warranted application.

The structure of the argument is as follows. Firstly, a principle called the Transparency Principle is developed. This places constraints on the individuation of concepts across differing contexts. The Transparency Principle is then used to show that Inferential Role Semantics is false because it cannot provide a satisfactory account of cognitive dynamics; that is, of the conditions under which a concept is retained through changes in the epistemic state of the subject over a period of time. A version of Conceptual Atomism is then defended and it is shown that this theory yields the correct individuation of concepts. According to this theory the concepts of an individual subject are individuated in terms of *referential episodes*, episodes of ongoing reference to an object or property during which it is diachronically transparent to the subject that the same thing is being referred to. The more general notion of a *referential practice* is then used to account for the sharing of concepts by more than one person. Finally, a novel account of the thoughts expressed using indexical terms is defended in order to show that indexicals present no counterexample to Conceptual Atomism. This account of indexical thoughts is of some consequence in its own right.

Introduction

Frege's idea is that being in the same epistemic state may require different things of us at different times; the changing circumstances force us to change in order to keep hold of a constant reference and a constant thought - we must run to keep still (Gareth Evans 1995, p. 308).

This thesis is about what follows from the application of a certain simple principle, which I shall call the Transparency Principle, to some matters in the philosophy of thought. The main conclusion that will be drawn (though by no means the only one) is that concepts, the components from which thoughts are composed, are unstructured entities an account of whose individuation makes no essential reference to other concepts in the possession of the thinking subject or to any particular means by which the reference of the concept is identified by the thinking subject. Following Jerry Fodor (1998) I shall call this general position *Conceptual Atomism* and the position it opposes *Inferential Role Semantics*. According to Inferential Role Semantics concepts are individuated by their inferential roles (that is, by the characteristic inferences which can be made using them) or their conditions of warranted application.¹ This puts them in constitutive relations to other concepts. Some versions of Inferential Role Semantics take the notion of 'inference' more broadly and, for example, allow rational relations to experiences of certain sorts to count as individuating; consequently advocates of this position argue that there are such things as 'recognitional' concepts.

¹ This definition is used informally by Fodor 2000, p. 350 to define 'epistemic individuation', which he contrasts with Conceptual Atomism. But note that this differs from how Peacocke (2000a, p. 300) uses the expression 'epistemic individuation'. Peacocke's use refers to concepts whose individuation involves conditions under which the subject may be said to have knowledge of a certain content: 'There is a class of concepts each member of which can be individuated in terms of the conditions for a thinker's *knowing* certain contents containing those concepts; and every concept is either such a concept, or is individuated ultimately in part by its relations to such concepts'. The class of concepts which Peacocke refers to as 'epistemically individuated' is therefore contained within the class to which Fodor refers; and if concepts were as Peacocke describes them they would all be epistemically individuated in Fodor's sense.

To make it clearer what Conceptual Atomism amounts to it is necessary to say more precisely how the word ‘concept’ is being used. I use the term in a way which has become standard in a certain area of the literature (Peacocke 1992, Fodor 1998) to mean a component of a thought, where the word ‘thought’ should be construed in a broadly Fregean way (I shall say more on this below). In his book *A Study of Concepts* (1992, p. 2) Christopher Peacocke gives the following stipulation about the way he uses the word ‘concept’:

Distinctness of concepts: Concepts C and D are distinct if and only if there are two complete propositional contents that differ at most in that one contains C substituted in one or more places for D , and one of which is potentially informative while the other is not.

Thus, for example, ‘ $C = C$ ’ is not informative but if ‘ $C = D$ ’ is informative then C and D are different concepts (I use the symbol ‘=’ to mean ‘is identical to’). Concepts can, on Peacocke’s definition, be of any category: singular, predicative or of higher level. Thus even names express concepts. I shall follow Peacocke in this. It should be noted that as well as differing from many everyday uses of the word ‘concept’, this is also completely different from the way Frege used the word (in translation) to mean the reference of a predicate, or a function from objects to truth values.

Peacocke’s stipulation more or less captures the intuitive notion behind the use of the word ‘concept’ as I shall use it. A concept, thus construed, is a component of a thought of a rational thinker. Peacocke’s criterion is clearly based on Frege’s criterion for distinguishing different thoughts, which Evans (1982, pp. 18-19) dubbed the Intuitive Criterion of Difference. This says that two sentences express different thoughts whenever it is rationally possible for a thinker who entertains both thoughts at the same time to assent to one but not to assent to the other. Peacocke has merely modified this in order to be able to individuate concepts rather than whole thoughts.

Although I shall differ from Peacocke by introducing an additional criterion which, in certain circumstances, determines when two concepts are the same I shall take it for granted that Peacocke and I, as well as other philosophers engaged in recent debates about concepts, are talking about the same thing. It therefore makes sense to disagree with what these philosophers say about the nature of concepts; we are not merely talking about different things.

I shall use the expressions 'sense' and 'mode of presentation' more or less as Frege used them but without any commitment to the traditional Fregean 'descriptive' view. Frege spoke of the sense of a sign, in which was contained the mode of presentation of the reference. Sense is thus a property of linguistic expressions whereas a mode of presentation is expressed by a term with a given sense but is itself a component of a thought. Thus when one uses a term with a certain sense this indicates that one is thinking of the reference under a certain mode of presentation. A thought is what is expressed by the sense of a complete sentence and is thus composed from the modes of presentation expressed by the terms which form the sentence. A concept and a mode of presentation are therefore almost the same thing; the only difference being that thinking of an object under a certain mode of presentation is the same thing as entertaining a certain concept. But for convenience I shall sometimes speak of entertaining a mode of presentation, which should be understood as the same thing as entertaining a concept.

It is crucial to the Fregean notion of a thought that thoughts can serve as the subject matter of a rationalising psychology. Thoughts are the entities to which we have attitudes such as beliefs and desires, and toward which a rational person cannot have conflicting attitudes. Frege also held that thoughts are truth evaluable, and I shall follow him in this as well. It is, of course, open to dispute (and indeed it has been disputed) whether there can be single entities which are both cognitively significant and truth evaluable simultaneously. One cannot make Fregean thoughts exist, or even make it possible that they could exist, just by giving a list of the properties which define them (just as one cannot, for example, legitimately speak of round squares just by defining

them as entities which are both round and square). For the purposes of argument, however, I shall take it for granted that there are such entities as Fregean thoughts. Even if this is incorrect, however, much of what I say about concepts will still be valid, though it would have to be understood in a slightly different way. But, in any case, at least one of the apparent obstacles to a Fregean theory (relating to indexical terms) will be removed by the account to be given. This does not show that a Fregean theory is obligatory but it at least suggests that such a theory can be made consistent.²

My adoption of this broadly Fregean framework should not, however, be taken to imply a commitment to Frege's own account of modes of presentation; on the contrary, this is precisely what is at issue. There is a certain very common and, it has to be admitted, highly plausible reading of Frege in which modes of presentation are seen as conditions which something must meet in order to be the reference of a term; whether or not this is what Frege really intended, my account will differ radically from it. I shall start by merely stipulating that modes of presentation are individuated in such a way as to capture the rationality of the thinking subject; then I shall consider what kind of entity could serve this purpose.

The Transparency Principle, the main tool which I shall use in investigating the nature of concepts, is a simple modification and generalisation of the Intuitive Criterion of Difference. The Transparency Principle yields the Intuitive Criterion of Difference as a special case. It allows an account to be given of the individuation of thought components (and hence of whole thoughts) in situations not covered by the Intuitive Criterion of Difference. In particular, it allows an account to be given of the circumstances under which a person can retain a concept through time and through changes of various sorts, and of the circumstances under which different people can

² See chapter 6. An important motivation for wishing to defend a Fregean conception of thought is the conviction that thoughts are representational intrinsically, not just contingently, and that there is consequently no gap between being in a given psychological state and being related to the world in a certain way. John McDowell (1998a, p. 243) has described the representational nature of thought as 'the most conspicuous phenomenological fact there is'; see his 'Singular Thought and the Extent of Inner Space' (1986) for a brilliant defence of what is, in effect, the Fregean conception of thought.

possess the same concept at a given time. The study of the former is known as *cognitive dynamics*.³

The central argument of the thesis will be that Inferential Role Semantics cannot give a satisfactory account of cognitive dynamics whereas Conceptual Atomism can (as well as being able to explain everything that Inferential Role Semantics can explain). It will also be shown by a further appeal to the Transparency Principle that Conceptual Atomism gives an intuitively more satisfying account of the conditions under which different people possess the same concept. These issues occupy the first of the two main parts of the thesis, consisting of chapters 2 to 4. The remaining part, consisting of chapters 5 and 6, deals with the special case of indexicals, which raise certain special difficulties for Conceptual Atomism. The resolution of these difficulties involves the development of a theory of indexical thought which should also be of some interest in its own right.

There is, I believe, nothing revolutionary about the Transparency Principle itself. It merely reflects a basic and familiar feature of thought and should therefore be seen as little more than a truism; something which would be very difficult to deny while still retaining a notion of thought recognisable as the subject matter of propositional attitude (folk) psychology. I try to make this clear while setting out the Transparency Principle in chapter 1. Then in chapter 2 I draw a distinction between *Ways of Thinking* and *Ways of Referring*. Ways of Thinking are just modes of presentation; I use the terminology only because it is already in use (mainly by philosophers influenced by the work of Gareth Evans) and because it helps in drawing the contrast with Ways of Referring. Ways of Referring to a thing, on the other hand, are different ways in which the thinker can identify the thing; they are different means by which the thinker can satisfy what Evans called *Russell's Principle*, the principle that in order to be referring to a particular thing the thinker must be able to distinguish that particular thing from all other things. By applying the Transparency Principle to certain dynamic

³ The expression 'cognitive dynamics' is due to David Kaplan (1989, p. 537).

cases I argue that, contrary to what has been assumed by a great many philosophers, Ways of Thinking and Ways of Referring are entirely separate things, and when a person thinks of a thing in a certain way this may be consistent with referring to it in any of a variety of different ways. This suggests a Conceptual Atomist view which, although highly Fregean and consistent with the views of neo-Fregeans such as Evans and John McDowell, also has a surprising amount in common with theories recently put forward by certain philosophers whose views have usually been thought of as antagonistic to the Fregean tradition, such as Jerry Fodor's recent work on concepts and David Kaplan's work on words. In chapter 3 the Conceptual Atomist view is expanded upon and I make a positive suggestion about how concepts should be individuated on an Atomistic theory. On this view, concepts are worldly entities which are individuated, like most worldly entities, by the circumstances of their creation and continuity through time; this leaves room for certain of their properties to change through time without a change in the identity of the concept. The consequences of this view for various issues including analyticity and holism are also discussed.

In chapter 4 the Transparency Principle is applied to the question of whether different people can share the same concept. The arguments are analogous to those of chapters 2 and 3. By applying the Transparency Principle to certain interpersonal (rather than dynamic) cases it is shown that despite the numerous objections which have been raised to it, the Fregean claim that communication involves the sharing of concepts is correct (though for reasons which will become apparent in chapter 6 I remain agnostic at this stage about whether communication always involves the sharing of thoughts). The objections to the Fregean view have two main sources. The first is a failure to distinguish Ways of Thinking from Ways of Referring. Because it is always possible for different speakers of the same language to employ different Ways of Referring when using the same linguistic term it has seemed obvious to the objectors that communication need not involve sharing the same Ways of Thinking either. This is easily seen as erroneous once the distinction between Ways of Thinking and Ways of

Referring is made. The second source of objections to the Fregean account of communication is a failure to take adequate account of the social character of language use. The objections tend to presuppose that each speaker uses words in a way which involves no commitment to participation in a communal word-using practice, as though each of us chooses our own vocabulary and it is only by a fortuitous coincidence that our words tend to match up. This is false. Once adequate account is taken of the commitments which have to be accepted whenever one acquires a term which belongs to a shared language it becomes much clearer why, at least under normal circumstances, communication involves sharing the very same concepts. The objection has, I believe, been encouraged by Davidson's view that radical interpretation begins at home, with its corollary that there is no such thing as a language. I argue that this is incorrect; Davidson is correct to argue that a certain kind of communication can take place without a shared language, but nevertheless there is another kind of communication which does involve a shared language and shared concepts.

In chapter 5 I argue in favour of the Fregean idea that when a thought is expressed using an indexical term the same concept can sometimes be retained through a change of context which requires the indexical term which expresses the concept to be exchanged for a different one. This is shown to follow from the Transparency Principle, though I examine a number of additional arguments in favour of the same conclusion. In chapter 6, however, I argue that the Fregean account is not adequate as it stands because it does not account for the changes in the inferences which may be drawn or the actions which it may be appropriate to perform when the indexical term changes. This is a well-known problem with regard to indexical communication and has been an additional reason why many philosophers have abandoned the Fregean account of communication, yet it seems to have been entirely neglected in discussions of cognitive dynamics. It poses a serious problem for Conceptual Atomism given that the latter claims that concepts do not have characteristic inferential roles. The answer cannot be to accept Inferential Role Semantics for indexicals, however, because

Inferential Role Semantics offers no adequate account of how the same concept can be retained through a change of indexicals, as argued in chapter 5. Fortunately these difficulties can be straightforwardly resolved by a novel account of indexical thought in which there is an *egocentric predicate* component associated with indexical thoughts. On this account a single, atomic concept can be retained through a change of indexicals while the associated egocentric predication changes according to the indexical term. Having described this theory in some detail, I argue that it is independently supported by certain hitherto neglected aspects of the phenomenon of immunity to error through misidentification.

Any theory of concepts is bound to have connections with a great many philosophical issues; too many for it to be possible to discuss them all in a thesis of this size. A number of subsidiary claims are, however, made along the way where it has been possible to do so reasonably briefly. Some of the more controversial ones have been placed in the three appendices and should be regarded as optional extras rather than as claims which are essential to the overall argument of the thesis.

Part 1: Preliminaries

CHAPTER 1

The Transparency Principle

The Intuitive Criterion of Difference - that two sentences express different thoughts whenever it is rationally possible for a thinker who entertains both thoughts at the same time to assent to one but not assent to the other - captures an important feature of the structure of rationality. Let us take as a necessary condition for being rational that a rational being will not knowingly have simultaneous conflicting attitudes toward the same content; 'knowingly', because it is always possible to have logically inconsistent beliefs without realising that they are inconsistent (simply failing to have worked through the logical consequences of all of one's beliefs to eliminate inconsistencies does not make one irrational). Now, there is nothing irrational about believing that Hesperus is bright while simultaneously believing that it is not the case that Phosphorus is bright, even if 'Hesperus' and 'Phosphorus' refer to the same thing, so long as the thinker is not aware that they co-refer. This is not due to a failure to work through the logical consequences of the beliefs, for there is no way for even a perfectly rational thinker to derive a contradiction from 'Hesperus is bright' and the negation of 'Phosphorus is bright' without the addition of further premises. This illustrates the well-known fact that the requirements of rationality cannot be captured solely at the level of reference and it is, of course, for this reason that Frege introduced the distinction between sense and reference. The level of senses and modes of presentation is intended to capture the rational bearing that one thought has on another. Hence Frege's Intuitive Criterion; whenever there does not exist the kind of rational relation between the thoughts expressed by two sentences that would require the thinker to hold the same attitude to both, the two thoughts cannot be the same.

The intuition behind Frege's criterion is what Michael Dummett (in numerous places, e.g. 1978, p. 131; 1991, p. 34) has called the claim that sense is 'transparent'; that is, that anyone who grasps the senses of two expressions must thereby know whether or not they are the same. In this chapter I shall argue that the notion of transparency can be used to give a satisfactory principle for determining the sameness of modes of presentation; and, moreover, one which will apply not only to the modes of presentation comprising the thoughts of a single thinker at a single moment in time but also to those comprising the thoughts of a single thinker at different times and the thoughts of different thinkers at the same time. This is the principle I shall call the Transparency Principle.

The aim, then, is to find out what constraints are placed on the individuation of modes of presentation by whatever rational relations hold between thoughts across different contexts. It is, of course, a substantive claim that there *are* any rational relations between thoughts which occur in different contexts, let alone that these relations can suffice to determine whether or not components of the thoughts are identical. The former claim will be defended as and when necessary for the particular variations of context with which we shall be concerned in later chapters, but it is first necessary to provide a more general principle to ground the general claim that the rational relations between thoughts can suffice to determine whether they contain the same modes of presentation. I shall first look at Frege's tentative attempts to give a criterion for sameness of thought which, like his Criterion of Difference, applies to a single speaker at a single time. This has been shown to be unsatisfactory as it stands, and was never published by Frege, but I shall argue that it can easily be modified to suffice for our current purposes (which do not require a completely general criterion for sameness of thought, but only one which is capable of settling the question of sameness of concept in certain situations).

1.1 *A criterion for sameness*

Frege stated his Intuitive Criterion of Difference, giving a sufficient condition for two sentences to express different thoughts, in numerous places. On just two occasions, however, he suggested a criterion for sameness of thoughts. Firstly, in a letter to Husserl (1980, pp. 70-71) he wrote:

It seems to me that an objective criterion is necessary for recognising a thought again as the same, for without it logical analysis is impossible. Now it seems to me that the only possible means of deciding whether proposition A expresses the same thought as proposition B is the following, and here I assume that neither of the two propositions contains a logically self-evident component part in its sense. If *both* the assumption that the content of A is false and that of B true *and* the assumption that the content of A is true and that of B false lead to a logical contradiction, and if this can be established without knowing whether the content of A or B is true or false, and without requiring other than purely logical laws for this purpose, then nothing can belong to the content of A, as far as it is capable of being judged true or false, which does not also belong to the content of B.

This criterion has been criticised (by Jean van Heijenoort (1977) and by Dummett (1981b); see also Picardi 1993) for the fact that it appears to make any pair of analytically equivalent sentences express identical thoughts, which presumably was not Frege's intention. The problem lies in the criterion's appeal to whether or not a contradiction can be established; in the same way that, as suggested above, one should not be deemed irrational for simultaneously holding beliefs from which a contradiction could be derived (given enough time and effort), by the same token two sentences should not be taken as expressing the same thought just because logical laws can be used to show that a contradiction can be derived from the assumption that they differ in truth value.

At around the same time, however, Frege also produced a slightly different formulation (which appears in his *Posthumous Writings* (1979), pp. 197-8):

Now two sentences A and B can stand in such a relation that anyone who recognises the content of A as true must straightaway also recognise that of B as true, and conversely, that anyone who accepts the content of B must immediately accept that of A (*equipollence*).... One has to separate off from the content of a sentence the part that alone can be accepted as true or rejected as false. I call this part the thought expressed by the sentence. It is the same in equipollent sentences of the kind given above.

By insisting that anyone who recognises the content (that is, grasps the sense) of the two sentences must recognise *immediately* whether they have the same truth value, the problem of mistaking analytic equivalence for sameness of sense is removed. This second version of the criterion, however, is still not satisfactory. Notice first that the criterion does not rely on sense being transparent if transparency is taken to imply that anyone who grasps the senses of two expressions must thereby know whether or not they have the same *sense* (as Dummett's (1991, p. 34) definition implies). Strictly speaking, what the criterion just quoted implies is that anyone who grasps the senses of two expressions must thereby know whether or not they have the same *reference*; this, in the case of complete sentences, being the truth value of the thoughts expressed. In some passages Frege seems to hint at the former notion of transparency (e.g. 1979, p. 211) and perhaps this would have suited his theory better given that he was prepared to allow there to be sense without reference. If there are empty names, for instance, one must have a means of distinguishing between their senses which makes no appeal to their references (though perhaps the notion of whether or not the thinker believed that they must share the same reference assuming that they refer might be of some use here).

Nevertheless, the stated criterion for sameness of thought is that whenever it is immediately apparent to the subject that two sentences must share the same truth value then the sentences both express the same thought; whenever, that is, it is transparent to the subject that the sentences must share the same reference. As Dummett (1991, chapter 14) points out, Frege seems to have had a criterion of this sort in mind when, in *Die Grundlagen der Arithmetik* (§64f), he appealed to the idea that the following pairs of sentences express the same ‘content’ as a means of introducing the concepts of direction and, by analogy, number:

(A1) a is parallel to b

(A2) The direction of a is the same as the direction of b

(B1) There are just as many Fs as Gs

(B2) The number of Fs is the same as the number of Gs

Frege’s idea seems to have been that we have an ‘intuition’ of the notion of parallelism which enables us to grasp the concept used in (A1), and from this it is possible to derive the new concept of direction (of which Frege claimed we have no such intuition) that occurs in (A2) because (A2) expresses just the same judgeable content as (A1). A similar move is made to introduce the concept of number using (B1) and (B2).

Given that *Die Grundlagen der Arithmetik* was written before Frege drew the distinction between sense and reference it would be unreasonable to presuppose his later notion of thought in criticising the move he makes here. These examples, however, provide a good illustration of why Frege’s second criterion of sameness is unsatisfactory. For the application of this criterion to these examples yields the conclusion that both of the sentences in each pair express the same thought; anyone who grasps their sense can immediately appreciate that they must share the same truth value. Yet it is intuitively clear that the sentences do not, in fact, express the same

thoughts at all. For thoughts with different components are different thoughts, and it is intuitively clear that the pairs of thoughts contain different concepts; indeed, that seems to have been Frege's whole point in discussing them.

The problem with Frege's criterion, then, is that there can be cases in which it is manifest to someone who understands two expressions that their reference (truth value, for whole sentences) must be the same without it being plausible that the two expressions share the same sense. It is implausible that the sense is the same because the two expressions differ in structure in a way which requires the subject to possess different concepts in order to be able to grasp them; and this is sufficient for them to have different senses.

Since the difficulties for Frege's criterion arise from differences in the structure of the sentences one solution would be to specify that two thoughts can only be identical if there are no differences in their structures. This would preclude the possibility that the same thought might be expressed by sentences whose surface grammatical structure differed, and this might seem too strong a requirement. It would, however, be possible to make a more modest claim: that when there are two sentences whose grammatical structure is the same, they express the same thought whenever it is immediately apparent to the subject that they must share the same truth value. This would not be a full criterion for the individuation of thought, but it would at least allow us to conclude that certain pairs of sentences express the same thought. The principle which I shall develop is of this more modest variety, but it will suffice for the purposes for which it is needed.

Before proceeding, however, it is necessary to mention a further potential difficulty with the idea of using transparency as a sufficient condition for sameness of mode of presentation.¹ There can be situations in which it seems transparent to the subject that two expressions have the same reference when in fact they do not. Consider, for

¹ The following remarks are much influenced by John Campbell's (1987) article 'Is Sense Transparent?'

example, two successive utterances of a demonstrative 'that'. It is possible to imagine (at least in principle) a situation in which the object referred to by the demonstrative is exchanged for a different one without the subject noticing. It would therefore seem clear to the subject that both utterances of 'that' had the same reference, yet this would not be correct. Similar problems can be encountered with the word 'here'. It would be possible to utter two sentences containing the word 'here' in quick succession without realising that one had moved during the time between the two utterances. The reference would have changed without this being apparent. Frege's own criterion applies only to thoughts which are entertained simultaneously, and therefore does not apply in this case. I shall, however, use the notion of transparency to argue (in chapter 5) that two successive utterances of a demonstrative can express the same mode of presentation, so the more general diachronic case must be taken into account here.

It is worth pausing to note that there is only a difficulty here if it is assumed that sense is intrinsically connected to reference. It is, as I implied in the introduction, a defining characteristic of the Fregean approach that the sense of a term is sufficient to uniquely determine its reference. There are those, however, who think that Frege should have kept the level of sense independent of the level of reference (e.g. Perry 1977), so that two people could entertain the same mode of presentation while thinking about different things. On such a view, it is trivially true that sense is transparent. The problems appear only when sense is understood as determining a unique reference. Nevertheless, this is the view I shall defend.

A simple response to the problem would be to hold that transparency can be used as a criterion for the individuation of modes of presentation only when the reference is the same. This requires abandoning a Cartesian conception of the mind according to which the mind is fully transparent to the subject and it is always apparent to the subject whether or not two thoughts are the same. Many neo-Fregeans have already accepted this conclusion (e.g. Campbell 1987, McDowell 1986). And as Campbell (1987) points out, although unnoticed switches of reference are possible in principle, the fact that it is

consequently possible to entertain a doubt about whether a demonstrative has preserved the reference of an earlier demonstrative does not suffice to show that the senses of two demonstratives are always different, rather than merely different when the reference is different. For we do not normally think in this reflective manner when keeping track of an object and thinking about it demonstratively over a period of time; there is not normally any question of making a reflective judgment about whether the object seen now is the same one that was thought about moments ago. One merely keeps track of the object without any need for such reflection, and in doing so one keeps hold of the thought.

While I think that this is correct, a further response to the reference switching problem is possible. The objection presupposes that each occurrence of a demonstrative has its reference determined independently of any other use of a demonstrative, so that there is no question that in the switching example the two uttered demonstratives (or the occurrences of 'here') have different references. But it seems to me that it is possible to doubt this, and to claim instead that the second demonstrative is, to use Evans's (1982, pp. 132-135) phrase, 'ill-grounded'. Suppose that the subject conceived of each demonstrative as an entirely independent occurrence, whose reference was determined purely by the object in view at that particular moment. Then, at the second utterance, it would be a matter for further investigation whether or not the same object was being referred to as had been referred to earlier. But, *ex hypothesi*, this is not the case; the subject takes it for granted on the occasion of the second utterance that there has been no change of reference.

The second use of 'that' must therefore be understood as linked to the first. The relationship seems to be quasi-anaphoric; the second 'that' is intended to preserve the reference of the first in a similar way to an anaphoric utterance of 'it', yet the dependence is only partial because the subject maintains an independent grip on the reference by keeping track of the object. The problem occurs when the two components, the quasi-anaphoric intention to preserve the reference of the first

demonstrative and the independent means of identifying the reference using perceptual information, come apart and pick out different objects. It then becomes unclear what the reference of the second demonstrative is; the thought is not well-grounded because the reference is ambiguous, and the demonstrative therefore lacks a clear sense, even though it will seem to the subject that the same sense has been preserved. The subject is simply in a state of confusion about the reference, without realising this. But if the same object had been referred to throughout there is no reason why the same sense should not have been preserved. It remains the case that transparency alone cannot without qualification be said to be sufficient for sameness of mode of presentation; but it does at least follow that where there is transparency there cannot be two different modes of presentation. Rather, in the case described, the subject starts by entertaining a certain thought but subsequently fails to retain it, and lapses into a confused state in which there is no clearly determined content.

1.2 The Transparency Principle

As I implied above, something less than a full-blown criterion for sameness of thought applicable to all situations will suffice for the investigations to follow. A criterion which covered thoughts expressed by sentences of arbitrary complexity might be technically quite difficult to formulate. For our purposes, however, we shall only need a criterion which determines the circumstances under which a mode of presentation is the same between two contexts. I take it for granted that there are primitive concepts from which more complex concepts can be constructed and whose properties determine the properties of the complex concepts constructed from them. Consequently we need only concern ourselves with these primitive concepts. In order to investigate these, it will be sufficient to concentrate on the very simplest sentences, atomic sentences of subject-predicate form.

For the purposes of argument we shall need to be able to assume that under certain circumstances two subject-predicate sentences share a component (either the subject or the predicate) in order to be able to consider whether the remaining component is also the same. This, however, is not problematic. Suppose we have two simple, atomic subject-predicate sentences ' a is F ' and ' b is F '. We need to be able to take it for granted that the ' $\text{is } F$ ' part is common to both in order to investigate whether ' a ' is the same mode of presentation as ' b '. Alternatively we might have the sentences ' a is F ' and ' a is G '; here we need to be able to take it for granted that the ' a ' part is common to both sentences, in order to investigate whether ' $\text{is } F$ ' and ' $\text{is } G$ ' are the same.

We can give a more general formulation of this as follows. Suppose $S(\gamma)$ is a function which forms an atomic subject-predicate sentence from an unstructured term ' γ '. For example, if ' γ ' is a singular term then $S(\gamma)$ might be the sentence ' γ is F ' where ' $\text{is } F$ ' is a predicate, or alternatively if γ is itself a predicate component then $S(\gamma)$ could be the sentence ' a is γ ', where a is a singular term. Then the question to be addressed is: if ' γ ' and ' μ ' are unstructured co-referential terms then under what circumstances does $S(\gamma)$ express the same thought as $S(\mu)$? In posing this question we take it for granted that the sentence-forming function ' S ' is the same in both sentences. But can we take this for granted without begging the question?

I do not think there is a genuine problem here. I shall be concerned, in what follows, with two kinds of case: those in which $S(\gamma)$ and $S(\mu)$ are uttered at different times by the same person, and those in which they are uttered at the same time by different people (perhaps 'uttered' should not be taken too literally here; I assume that an utterance can express the content of a thought, but the thought can still be present without any actual utterance. The sentences in question are those that the subjects would utter in order to express the thoughts, if they were to express them). Now, in order to investigate the circumstances under which ' γ ' and ' μ ', in the sentences $S(\gamma)$ and $S(\mu)$, express the same mode of presentation it will not be necessary to have a

separate criterion for whether or not the same sentence forming function 'S' appears in both cases. We can merely stipulate that it does. In order to be justified in making this stipulation we have only to assume that it is *possible* for 'S' to appear twice; we need not apply any criterion for establishing whether this is the case for any given pair of sentences.

The assumption that it is possible for 'S' to appear twice amounts to nothing more than the assumption that the same concepts can occur in different contexts; that is, they can be entertained at different times and by different people. To deny this would be to adopt a view of concepts as 'atoms', only capable of existing at a single moment in time in the mind of a single subject. This is particularly problematic in the temporal case. The word 'green' as it appears in 'grass is green' would have to express a different concept from that expressed by 'green' in 'emeralds are green', or even in two separate utterances of 'grass is green'. It is hard to see what could motivate such a view of concepts. If concepts were to be individuated in this way an entirely separate account would have to be given of the rational relations between one thought and another. It can be inferred, for instance, from the two thoughts just mentioned that both grass and emeralds have the same colour. It seems natural to account for this by appeal to the fact that the same concept, 'green', occurs in each thought. If this cannot be assumed then some explanation has to be given of how one occurrence of 'green' bears upon another. One might, of course, have all sorts of reasons for dividing up whatever one calls 'concepts' in various ways according to the project by which one is motivated. But the project engaged in here is that of giving a theory of thought in such a way as to make rational sense of a thinker. Given that this is the project, the atomic view is ruled out. Only a view on which the same concepts can occur on different occasions, either in the same thought or in different thoughts, can possibly hope to capture the structure of rationality.²

² See Evans 1985, pp. 309-311 for a related argument, and see Evans 1982, pp. 235-236 on the relation between persistence of beliefs and the requirements of rationality.

These considerations make it clear that it must sometimes be possible for a person to entertain thoughts containing the same concept at different times. But it may appear that there is still room for a theory in which different people possess entirely different concepts from one another. Rationality is, after all, an intrapersonal rather than an interpersonal matter; it is irrational to knowingly contradict oneself but it is not irrational to knowingly disagree with other people. It may not seem obvious, therefore, that there are any rational relations between the thoughts of different people which require appeal to shared concepts. Hence, although it is clear that a single subject may be assumed to entertain thoughts $S(\gamma)$ and $S(\mu)$ where 'S' is common to both, it may appear unjustified to take it for granted that these two thoughts could be entertained by different people, while still assuming that 'S' is common to both.

A certain amount of 'bootstrapping' may therefore be necessary in the interpersonal case, but I think this can be justified. If, as I shall argue in chapter 4, there can in fact be the relevant kinds of rational relations between the thoughts of different thinkers even though rationality is an intrapersonal matter, then some explanation of these rational relations must be given. I shall argue that if it is assumed, under certain circumstances, that there is a common thought component 'S' between the thoughts, $S(\gamma)$ and $S(\mu)$, of two different thinkers, a satisfying explanation of the rational relations between the thoughts can then be given, but only if ' γ ' and ' μ ' express the same concept. This gives an entirely consistent theory. If, on the other hand, it is not assumed that there is a common thought component 'S' then the rational relations between $S(\gamma)$ and $S(\mu)$ would be left entirely mysterious. This justifies the assumption that 'S' can be shared.

The foregoing threads can now be pulled together in the form of a general, if somewhat abstract, principle which, I believe, embodies the basic Fregean insight about the relation of transparency to the individuation of modes of presentation:

The Transparency Principle: If an attitude A to the thought expressed by the sentence $S(\gamma)$ can make it irrational to dissent from attitude A to the thought expressed by the sentence $S(\mu)$, without the involvement of any further thoughts, then ' γ ' and ' μ ' express the same mode of presentation.

An 'attitude' just means belief, desire or any other attitude which someone may have towards a thought. The qualification 'without the involvement of any further thoughts' is to allow for the possibility that it might be sometimes be irrational to take conflicting attitudes to $S(\gamma)$ and $S(\mu)$ because of some further thought such as a belief that $\gamma = \mu$. The principle should be understood as applicable only when there is no thought whose absence would allow conflicting attitudes to be taken to $S(\gamma)$ and $S(\mu)$.

It may help clarify things if we consider a familiar example in which the attitude 'A' is belief and $S(\gamma)$ and $S(\mu)$ are the thoughts of a single thinker at a single time. Suppose, then, that ' γ ' and ' μ ' are 'Hesperus' and 'Phosphorus' respectively, and that $S(\gamma)$ and $S(\mu)$ are the sentences 'Hesperus is bright' and 'Phosphorus is bright'. What the Transparency Principle says is that if believing that Hesperus is bright can make it irrational to disbelieve that Phosphorus is bright, without the involvement of a further belief such as 'Hesperus = Phosphorus', then 'Hesperus' and 'Phosphorus' express the same mode of presentation (the reason for the qualification '*can* make it irrational' will become apparent shortly). As a matter of fact it is possible to take conflicting attitudes to 'Hesperus'- and 'Phosphorus'-sentences so 'Hesperus' and 'Phosphorus' do not express the same mode of presentation. Whereas the Intuitive Criterion of Difference embodies the notion that where there can be conflicting attitudes there must be different content, the Transparency Principle uses a similar intuition to show that where there cannot be conflicting attitudes there must be the same content.

I have deliberately formulated the Transparency Principle in such a way as to leave it open whether $S(\gamma)$ and $S(\mu)$ are uttered at the same time or by the same speaker.

Formulating it in this way does not presuppose that there are rational relations between thoughts in different contexts. The existence of such rational relations must be argued for in particular cases. But whenever those rational relations exist, whenever it can be rationally impossible to take conflicting attitudes to $S(\gamma)$ and $S(\mu)$ without the involvement of any further thoughts then ' γ ' and ' μ ' express the same mode of presentation.

Let us now consider how the Transparency Principle applies to a diachronic case, one in which $S(\gamma)$ and $S(\mu)$ are uttered by a single speaker but at different times. The principle implies that if assenting to 'Hesperus is bright' at time t_1 can make it irrational to dissent from 'Hesperus is bright' at a later time t_2 without the involvement of any further thoughts then 'Hesperus' expresses the same mode of presentation on both occasions. The rational relation between the different times is such that what it is rational to assent to at a given moment can be systematically dependent on what it was rational to assent to at an earlier time. In simple cases like the example just given, the most straightforward way of capturing this systematicity is to regard the later thought as simply a continuation of the earlier one. As an objection to this one could, of course, claim that it is just a brute fact that certain modes of presentation stand in certain rational relations to other modes of presentation under certain circumstances. But in order to make such a claim it would be necessary to explain what could motivate one to regard the two modes of presentation as different, as well as explaining why the rational relations exist between some modes of presentation but not others. Given that the project being undertaken is to individuate thoughts in such a way as to capture the rational structure of thinking, it is very hard to see what motivation there could be. I shall give more detailed arguments of this general form when we come to deal with specific cases.³

³ Jérôme Dokic, at least in Dokic 1996, dissents from the view that diachronic transparency can be used as a criterion for sameness of thought (though in conversation he has expressed possible subsequent doubts). Dokic's strategy is to argue as follows: when there are two *disconnected* episodes

It is not, of course, irrational to change one's mind, and to take a different attitude to a thought from the attitude that one had earlier. Hence it would be rationally possible to take conflicting attitudes to $S(\gamma)$ and $S(\mu)$ at different times even if ' γ ' and ' μ ' expressed the same mode of presentation. One could start by believing $S(\gamma)$ at t_1 but by t_2 one's attitude could have changed in such a way as to make one dissent from $S(\mu)$. It is for this reason that I have formulated the Transparency Principle so as to require only that an attitude A to the thought expressed by $S(\gamma)$ *can* make it irrational for a thinker to dissent from the same attitude to $S(\mu)$. This should be understood to mean that there are imaginable circumstances, those in which no change of mind has occurred, in which the rational relation in question would hold.

Similar remarks can be made about the interpersonal case, in which $S(\gamma)$ and $S(\mu)$ are uttered by different speakers. As remarked above, there is nothing irrational about dissenting from the thoughts of others. If, however, it can be shown that there are circumstances in which one can take a certain attitude to someone else's thought, and that taking this attitude makes it irrational to dissent from that attitude to a certain thought of one's own, it would follow that a mode of presentation was shared by both speakers.

There is one more general issue to be dealt with. I have suggested a way in which the question of whether two terms express the same mode of presentation can be settled by appeal to the rational relations between the thoughts containing them. Now,

of thinking at different times it is an open question whether the later thought is the same as the earlier one. Because of the disconnection (i.e. there is a pause between the thought episodes) there is no diachronic transparency. But Dokic infers from a principle of ontological parsimony that, for instance, in a case in which a demonstrative thought is followed by a memory, then a subsequent demonstrative thought about the same object, since the demonstrative thoughts should be regarded as the same the mode of presentation held in memory after the initial perceptual encounter must be different from the initial demonstrative sense. I think that this is incorrect because it neglects the rational relations between the original demonstrative and the thought retained in memory, and I believe that this rational relation should take precedence over principles of parsimony. In any case, as I hope will become apparent, the principle that one should not multiply entities beyond necessity favours my theory because it avoids populating time with endless sequences of different modes of presentation. Arguments such as Dokic's strike me as examples of what happens when one loses one's grip on the

there may appear to be a problem concerning the direction of explanation implied by this. If the presence of a certain rational relation were understood as being what constituted the sameness of the modes of presentation then the Transparency Principle would be true, but trivially so. It would amount to nothing more than a description of the rational relations, rather than an explanation. Coupled with this there might be a suspicion of the notion of a mode of presentation implied by the exclusive use of the Transparency Principle as a means of individuating modes of presentation. If nothing else can individuate modes of presentation then they appear as somewhat mysterious, ineffable entities. At the very least, a theory in which thoughts are individuated in this way seems incompatible with naturalism, insofar as naturalism is construed as requiring that a theory of intentionality should be capable of being given in non-intentional terms (though there are, of course, versions of naturalism which do not require this).⁴

The worry is misplaced, however, because the direction of explanation just described is not the intended one. The Transparency Principle should be regarded, like Frege's intuitive criterion of difference, as a condition of adequacy on a theory of concepts, not as a theory of concepts itself.⁵ It provides, in a manner of speaking, the data which a theory of concepts has to explain; a way of mapping the rational relations between thoughts which any satisfactory theory of concepts must yield.

point of talking about modes of presentation which is, as I have been emphasising, to capture the rational structure of thinking.

⁴ Robert Stalnaker (1984, pp. 24-5) expresses a worry of this kind.

⁵ Christopher Peacocke (1986, p. 5) makes a very similar point about the use to which he puts 'Frege's Principle', the principle that 'if it possible rationally to believe that p and not to believe that q , then the contents that p and that q are distinct'. I shall argue in the next chapter that although Peacocke's (1992) theory of concepts conforms to Frege's Principle it fails to conform to the Transparency Principle.

Part 2: Concepts

CHAPTER 2

Concepts I: What Concepts Cannot Be

The main purpose of this chapter is to show that the main opponents of Conceptual Atomism cannot give a satisfactory account of cognitive dynamics. I shall give a more detailed characterisation of their views and the way in which they differ from Conceptual Atomism before setting out the argument against them, but first it is necessary to introduce some terminology.

2.1 Ways of Thinking and Ways of Referring

The expression 'Way of Thinking' has, as a result of Evans's (1982) work, gained some currency as an alternative expression for 'mode of presentation'. Evans introduced it to emphasise the fact that, according to him, a mode of presentation is a way of thinking *of* something and as such cannot exist unless there exists a certain thing which is being thought of. I shall make use of this expression both for the same reason as Evans and because it will be helpful to draw a contrast between Ways of Thinking and what I shall call 'Ways of Referring'.

A Way of Referring has something in common with what Frege called a 'mode of determination' (Beaney 1997, pp. 65, 321; but see qualifications below); it is the way in which the reference of a term is identified by the thinker. I shall take it for granted that in order to make a judgment about something it is necessary for the thinker to know which item is being referred to, which implies having some means by which to distinguish it from all other things. Russell (1912) argued for this view, and Evans (1982, chapter 4) defended it at length against theories which assume that reference

can be determined by something less substantial such as the causal history of the term. I shall follow Evans in calling it 'Russell's Principle'. I shall take it for granted that Russell's Principle applies to terms of all categories, not just singular terms. Thus in order to possess a concept of any category it is necessary to have a way of identifying whatever the concept refers to; that is to say, it is necessary to have a Way of Referring to it.

Although I shall not give a detailed defence of Russell's Principle here, I should nevertheless like to briefly note in passing a difficulty that would be encountered by its wholesale rejection. If the principle is not adopted in some form then this would leave open the possibility that someone could be said to be thinking about the world around them despite the fact that they possessed no means of identifying the objects or properties that they were supposed to be thinking about. Such a person would be unable to manifest their alleged thoughts in any way; they would be unable to act on any thought because they would be unable to establish that any encountered object was in fact the object of the thought. The fact that the alleged mental states were, as a matter of fact, caused by interaction with the object in question would make no difference to this. So the rejection of Russell's Principle would make possible only an extremely dubious 'private' conception of thought according to which it is possible for someone to be thinking about an object even when there are no circumstances in which an observer could discover this. Few philosophers would be willing to defend such a view; most, I think, would prefer to defend a conception of thought according to which thinking about x implies, at least in principle, being capable of acting on x in some way (with suitable qualifications to deal with objects only existing in the distant past etc.)

Ways of Referring fall into three main categories. These involve either perceptual contact, recognitional capacities or definite descriptions. I shall refer to these as

demonstrative, recognitional and descriptive Ways of Referring respectively.¹ Sometimes a Way of Referring may involve a mixture of different categories. I shall also speak of 'identifications' (as in 'a demonstrative identification of an object'), because a Way of Referring to an object is, as discussed above, a way of identifying it or distinguishing it from all other things. It will facilitate a lucid exposition of the arguments to follow if I make clear some assumptions that I shall make about how the different varieties of reference function before proceeding. My account of this is strongly influenced by Evans's book *The Varieties of Reference* (1982).

Demonstrative Ways of Referring to objects allow the subject to distinguish an object from all other objects by receiving a continuing stream of perceptual information from it. This does two things. Firstly, it determines which object it is whose states are treated as relevant to the truth or falsity of the relevant thoughts. Secondly, it determines which object will be acted upon in relevant situations. This occurs because the ongoing information link to the object allows the subject to 'home in' on its location and act upon it. This is the case even when the object is not where it appears to be because of a perceptual distortion of some kind, for example when an object is seen at the bottom of a swimming pool which looks shallower than it really is. Although the object is not where it appears to be, the ongoing information link allows the subject to home in on it and act upon it. These are the things which determine which object is the reference of the thought.²

Demonstrative identifications of properties are a little less straightforward. They do not seem to occur in practice anywhere near as often as demonstrative identifications of objects. Demonstrative identifications of properties often involve demonstrative identifications of objects. For example, it is possible to pick out 'that shape' in virtue

¹ See Evans 1982, p. 89 on this trichotomy. As Evans notes, it is also found in the works of Strawson and Dummett.

² Evans is sometimes interpreted as having held that what is crucial to a demonstrative Way of Referring is that it enables the subject to 'locate' the object egocentrically. See for example McDowell 1990. However Peacocke (1991) is surely correct to argue against McDowell that being able to 'locate'

of a demonstrative identification of an object which has the shape in question. The shape might, for example, be sufficiently complex or irregular that the subject could not remember it well enough to distinguish it from a slightly different shape on a subsequent occasion. In such a case, the subject's ability to identify the shape rests entirely on an ongoing ability to identify the object which has that shape. It is not clear, however, that the demonstrative identification of a property always rests on an identification of the object whose property it is. It is, for example, possible to perceive properties relating to the way something sounds without being able to identify the source of the sound. One could then refer to 'that sound' (meaning 'that *kind* of sound') without identifying its source. Sometimes one's identification of the sound would not rest entirely on one's ongoing perception of the sound; one might be able to recognise the sound on a subsequent occasion and think 'there is that sound again'. But there might be sounds of a particularly unusual nature such that one could not be sure that it was the same sound again on a subsequent occasion. In such a case one's ability to identify the sound would rest entirely on one's ongoing perception of the sound, but this would not necessarily involve the perception of an object.³

It is a matter for debate whether, and to what extent, demonstrative identifications must be supplemented with sortal terms. It could be argued that it is impossible to truly identify a thing without being aware that the thing falls into some very general category such as 'physical object'.⁴ If this is correct then it indicates that demonstrative identification always involves at least a minimal recognitional element, wherein the

the object should be understood to imply being able to 'home in' on the object even if it is not quite where it seems to be. This seems correct both in itself and as an interpretation of Evans.

³ Note that during an episode of perception it might be possible to detect small changes in the sound which it would not be possible to detect if the change occurred after a break in perception. Otherwise, if it were not possible to detect such changes, then it would be possible to argue that the perception of the sound does nothing more than 'fix the reference', setting up a recognitional capacity (a disposition to recognise the same sound on subsequent occasions). It would then be the recognitional capacity which would enable the sound to be distinguished from all others, rather than the perception itself.

⁴ See Recanati 1993, pp. 169-72, for an argument that no 'specific' sortal is required for demonstrative identification (as specific as, for example, 'bird' or 'plane') but that there may still have to be some very general category such as 'perceptible and occupying a certain portion of space'.

object is partially identified in virtue of some (very general) property which it possesses.

Recognitional Ways of Referring are intrinsically related to perceptual encounters but enable objects (or properties) to be thought about while not being perceived. The expression 'recognitional capacity' should be understood as meaning a disposition to recognise a perceptually presented object or property unreflectively. The capacity to recognise human faces is a good example; one often simply sees that it is or is not so-and-so without needing to reflect upon what one is looking at. This contrasts with partially descriptive cases in which one can find oneself thinking that the person looks a bit like so-and-so but being unsure whether it is really so-and-so because one had the impression that so-and-so was a bit taller and wore glasses. One thus reflects upon whether the perceived person satisfied certain descriptions.

Recognitional capacities often have to be supplemented with additional criteria or information before a unique reference can be identified. There can be no guarantee that a given object does not have a duplicate, so merely recognising the appearance of something is sometimes not sufficient to uniquely identify it. Consider the following example (taken from Evans 1982, pp. 279-80). Suppose that I am able to see a sheep on a hillside where I am sitting. There may be certain features of the sheep which would make it easy for me to recognise it if I took my eye off it for a while (if it is the only sheep around then I just have to be able to recognise a sheep). In general, however, there will be other sheep somewhere in the world which I would not be able to distinguish from the present one by appearance alone. If there is only one sheep on this side of the valley, however, then given certain assumptions about the maximum speed at which sheep can travel when I close my eyes for a short period of time I can be confident that the sheep I see when I open them is still the same one. The sheep could not have left the area and no other sheep could have reached the vicinity during

Something slightly more substantial than this may be required in order to keep track of a persisting object as it moves around (*ibid.*, p. 171n).

the time that I had my eyes closed. So in cases like this one my visual recognitional capacity is supplemented by a constraint on the spatiotemporal region in which the reference must be found. The identification of this spatiotemporal region will typically be demonstrative.

There is a group of concepts whose reference can only be identified via a kind of recognitional capacity, which may or may not be purely perceptual. These are 'mind-dependent' concepts: concepts whose references are determined solely by the ability of human beings to recognise them or judge that they instantiate the concept. Jerry Fodor (1998, chapter 6) has recently given a detailed and highly illuminating discussion of these concepts.⁵ I shall not reproduce the details of Fodor's position here but, briefly, the idea is that a mind-dependent concept such as DOORKNOB refers to those items which we are disposed to call 'doorknobs' and there is no other way of individuating the class of items referred to than through the fact that we have such dispositions.⁶ There may be no way to define a doorknob in terms of its form or function. It has rather to do with the role that certain items play in the lives of human beings within certain cultures. A determinate, objective reference is nevertheless determined for the concept. Doorknobs may be sometimes be recognised by perception alone, but since it would also be possible to imagine a novel object and decide that it is a doorknob without having perceived it, and since recognising a doorknob may involve thinking about aspects of its function that are not immediately perceptible, then strictly speaking it is not purely a perceptual recognitional capacity which determines the reference.

We have already seen how descriptive elements can come into play in otherwise demonstrative or recognitional identifications. But it is also possible to give a unique identification of an object or property using a definite description. Sometimes the use of a descriptive identification will rely on a demonstrative or recognitional

⁵ The idea of a mind-dependent concept bears some relation to the view of colour concepts put forward by John Campbell (1993) and the view of value concepts put forward by David Wiggins (1987, chapter 5).

⁶ Following a common convention I use words in capital letters for the names of concepts.

identification of something referred to by one of the primitive concepts from which the description is composed. For example suppose that I can identify something as ‘the only black sheep on this side of the valley’. This picks out a unique individual but its ability to do so rests on my being able to identify the region of space that I refer to as ‘this side of the valley’. This will probably rest ultimately on a demonstrative identification of some region of space. I shall not assume, however, that descriptive Ways of Referring always involve a *guarantee* that there is a unique reference. For any definite description ‘the ϕ ’ which does not contain any demonstrative elements it is always impossible to rule out the possibility that there is more or less than one item which fits the description.⁷ It is often assumed as a consequence of this that the description requires a spatiotemporal ‘anchoring’ using a demonstrative to ensure a unique reference, as in the example just described. By contrast, however, I shall assume that if there is, as a matter of fact, one and only one ϕ then the description ‘the ϕ ’ could be a Way of Referring. This may be a rather controversial assumption but fortunately very little of what follows depends on it; those who disagree can simply assume that whenever I speak of a descriptive Way of Referring there is always a demonstrative anchoring present within the description. I do not, however, assume that there could be a descriptive Way of Referring in which the elements of the description involve neither a demonstrative nor a recognitional Way of Referring. Such a description would be empty; it could never be used to identify anything in relation to either perception or action.⁸

Now, there are two ways in which the notion of a Way of Referring could be understood; either it does, or it does not presuppose that there is something referred

⁷ Cf. the possibility of what Strawson (1959) calls ‘massive reduplication’.

⁸ I give my reasons for rejecting the need for demonstrative anchoring in appendix B. I also discuss the difference between genuine singular thoughts (whose truth conditions involve a specific individual) and thoughts which are logically descriptive (whose truth conditions involve whichever object happens to satisfy a certain description) in appendix A. In addition to the cases discussed so far there are also cases in which in order to identify something it is necessary to defer to other people. Demonstrative identifications may ultimately be involved in these cases as well (for instance in being able to identify the people to whom I defer). I discuss deference in greater depth in chapter 4.

to. A Way of Referring which can exist without actually referring to anything corresponds to what Frege called a 'mode of determination'; it is a 'route to a reference', a condition which something must meet to be the reference or, more generally, a means by which a reference can be determined. The same means can be used to determine different objects in different contexts or no object at all. This contrasts with the use that Frege sometimes (but not always) made of the expression 'mode of presentation', wherein a mode of presentation essentially refers and can therefore be construed (on one interpretation of Frege) as involving a reference in addition to a mode of determination.⁹ Unless Ways of Referring are assumed to be identical to Ways of Thinking, however, there is nothing to prevent us from following Evans in holding that a Way of Thinking always has a reference while construing the notion of a Way of Referring as equivalent to a Fregean mode of determination, which has no essential reference. This is in fact the way in which I shall construe Ways of Referring.

2.2 Ways of Referring and the individuation of concepts

The difference between Conceptual Atomism and the views to which it stands opposed can be put in terms of the relation that the different views imply between Ways of Thinking and Ways of Referring. Conceptual Atomism holds that concepts are unstructured entities an account of whose individuation makes no essential reference to other concepts in the possession of the thinking subject or to any particular means by which the reference of the concept is identified by the thinking subject. Let us consider the two clauses of this definition in turn. Firstly, an account of the individuation of a given concept makes no essential reference to other concepts in the possession of the

⁹ See Beaney 1997, p. 23, for a description of Frege's relative uses of 'mode of determination' and 'mode of presentation'.

thinking subject. This rules out versions of Inferential Role Semantics according to which a concept is constitutively associated with a particular descriptive Way of Referring. Consider, for example, the claim that in order for someone to possess the concept HESPERUS they must identify its reference using the description 'the morning star'. It would follow from this that anyone who possesses that concept must be prepared to make certain inferences; if, for example, they believe that:

(1) Hesperus is bright

then they must be prepared to infer that:

(2) There is a bright star in the morning

Hence, it is said, the concept is individuated by its characteristic inferential role.

The second clause of the definition states that an account of the individuation of a given concept makes no essential reference to any particular means by which the reference of the concept is identified by the thinking subject. This simply generalises the first clause to cover other, non-descriptive Ways of Referring. Suppose that it were claimed, for example, that to possess a particular concept SQUARE it is necessary to be able to identify square objects when they are visually presented in an orientation which makes certain symmetries salient to the subject. This would mean that the subject would associate a certain recognitional Way of Referring with the concept. It is not clear that such a person would make any characteristic inferences using the concept of the kind supposedly associated with HESPERUS in the example above. But it might nevertheless be claimed that when the subject is visually presented with a square object in the appropriate orientation the subject has a reason to make certain judgments such as 'that is a square'. Thus the concept is individuated by its characteristic role in

judgments; and views of this kind are sometimes described as versions of inferential Role Semantics.

The essential difference between Conceptual Atomism and Inferential Role Semantics is therefore that according to Inferential Role Semantics an account of the individuation of a concept (and thus of a Way of Thinking) makes essential reference to a Way of Referring associated with the concept. Conceptual Atomism denies this. According to Conceptual Atomism, for example, there is no such thing as a 'demonstrative concept' or a 'recognitional concept'; there are just concepts, whose individuation does not depend on the way in which the reference is identified.

It may strike the reader that I have taken it for granted that Inferential Role Semantics claims that the characteristic inferential role of a concept relates to the way in which the reference is *identified*. Would it not be possible to put forward a different kind of theory according to which an account of the individuation of a concept makes no essential reference to a Way of Referring but nevertheless makes essential reference to other concepts in the possession of the thinking subject? It would be possible to claim, for example, that in order to possess the concept HESPERUS it is necessary to believe that Hesperus is the ninth planet from the sun, even though this is false (I am not suggesting that such an account would be at all plausible). Some other account would then have to be given of what determines the reference of HESPERUS.¹⁰

As a matter of fact, the arguments to be given against Inferential Role Semantics could be very simply modified to refute theories of this kind. I shall not say much about this, however, since as far as I know no such account has ever been put forward. The reason for this is simple. The key idea behind Inferential Role Semantics is that what

¹⁰ A well-known example from the literature which could be used to illustrate this point is Kripke's (1980, pp. 83-92) 'Gödel-Schmidt' argument. Kripke invites us to imagine that someone has only one identifying belief about Gödel (that he proved the incompleteness of arithmetic) and that this belief is false, and actually identifies Schmidt. But on a 'causal' theory the reference of 'Gödel' is Gödel, not Schmidt, because of the chain of causation going back from the use of 'Gödel' to the 'initial baptism' at which Gödel was named. In principle someone could hold that the mode of presentation expressed by the word 'Gödel' for this particular person is partially constituted in terms of the description 'the

gives a concept its identity is the characteristic role it plays in the judgments that the thinker makes, whether this be cashed out in terms of judgments based on inferences, on perception, or both. The thought is then that the role that a concept plays in judgments depends on the way in which the reference of the concept is identified. Other kinds of relations to other concepts would contribute nothing positive to the judgments that the thinker makes, given that judgment essentially aims at truth.

In fact Conceptual Atomism, at least in the version that I shall defend, has in common with Inferential Role Semantics the claim that the possession a concept requires the ability to identify the reference of the concept. Where the two views differ is in the fact that according to Conceptual Atomism nothing further needs to be added regarding the way in which this is achieved; for any given concept there must be a Way of Referring but this need not be any *particular* Way of Referring and the way in which the reference is identified plays no essential part in an account of the individuation of the concept.

2.3 Varieties of Inferential Role Semantics

I shall now give some specific examples of versions of Inferential Role Semantics. This will not be an exhaustive list by any means; for there seem to be almost as many views about what a concept is as there are philosophers. Instead I shall just mention a few well-known ones to give a more concrete idea of which views are in question. Many of these views have their roots in a traditional understanding of the idea that Fregean sense ‘determines’ reference. On this view a mode of presentation is seen as a means by which the reference can be identified or a criterion which something must meet in order to be the reference of a term; in other words, a Way of Referring. There are

man who proved the incompleteness of arithmetic’, even though this description does not pick out the reference of the concept.

certainly many passages in which Frege appears to suggest that wherever there is a difference in a Way of Referring there is a difference in sense:

An object can be determined in different ways, and every one of these ways of determining it can give rise to a special name, and these different names then have different senses; for it is not self-evident that it is the same object which is being determined in different ways. ('Letter to Jourdain, Jan. 1914', Beaney 1997, p. 321).

When Frege speaks of the way in which an object is 'determined' he means the Way of Referring to that object. I shall not engage in the dispute over the correct interpretation of Frege here. But there is clearly an interpretation of this passage according to which it is being argued that concepts must be individuated by Ways of Referring because if the object is referred to in a different way it may not be apparent to the subject that it is the same object. A difference in the Way of Referring therefore corresponds to a cognitive difference, and if two concepts differ in their cognitive significance they are different concepts. Consequently it is inferred that concepts are individuated by the fact that each one is associated with some particular Way of Referring.

That seems to be the basic idea that lies behind quite a few views but it appears in many different guises.¹¹ Most philosophers these days would reject a view which says that some or all concepts can simply be *defined* in terms of other concepts. Philosophical difficulties with definitions are well known. Ever since Quine published *Two Dogmas of Empiricism* (1961) the notion of analyticity that goes hand-in-hand

¹¹ One version of Inferential Role Semantics about which I shall say very little is the one advocated by Robert Brandom (1994). Brandom's position is very complex and consequently I cannot discuss it in detail but his 'inferentialism' (according to which concepts are individuated by their inferential roles, their representational properties being a product of this individuation rather than an element in it) is straightforwardly susceptible to the criticisms of Inferential Role Semantics given below. But this should not be taken to imply that I advocate what Brandom would call 'representationalism', according to which it is only an afterthought that conceptual contents are involved in inferences. I agree with McDowell (1997) that there is room for a position in between Brandom's dichotomy of inferentialism and representationalism according to which it is not merely an afterthought that concepts are involved in inferences but nevertheless their representational function is an essential part of what individuates them.

with the notion of definition has been treated with suspicion. Even so, there are still many adherents to the notion that there is a small subset of the overall set of concepts which can be defined analytically; for example by defining BACHELOR in terms of unmarried men.¹² And it also seems quite common to assume that even if concepts do not stand in analytic relations to other concepts there are relations between them which are *held* as defining by a given subject. This allows a view of concepts according to which the relations to other concepts in which a given concept is held determine which inferences involving that concept the subject will hold to be valid. A combination of the rejection of the analytic/synthetic distinction with an adherence to the assumption that concepts are individuated by conceptual relations has led some people to endorse a kind of radical holism in which there is no distinction between defining beliefs about something and informative beliefs about it, and any difference whatsoever in what is believed about X thereby makes X a different concept. Others have been more restrained, and held that there are only 'local holisms', groups of concepts which one cannot possess without possessing the other members of the group. But this has still been taken to imply that the concepts are individuated by their relations to other members of the group.

A more sophisticated view due to Christopher Peacocke (1986) says that concepts are individuated in terms of their canonical acceptance conditions, which are spelled out in terms of their 'canonical grounds' and 'canonical commitments'. Canonical commitments, for example, are those things which a thinker is committed to in judging that a given concept applies. In judging that something is cubic, for example, one takes on commitments about how the object would look from a different angle. If the object were to fail to fulfil these conditions one would be obliged to withdraw the judgment. The underlying assumption is still the same, however; that the means by which the reference is determined is partially constitutive of the concept:

¹² My own account of cases in which there is a strong intuition of analyticity is given in section 3.5.

If the truth condition for a content obtains, the spectrum of commitments for that content are fulfilled. This is quite consistent with the nonexistence of a route back from reference to sense: for it can be informative to say of two spectra of commitments that one is fulfilled if and only if the other is. So there can still be a unique content-determining association of a spectrum of canonical commitments with a given content. Thus suppose we could both feel heat and see it with infra-red vision. If the terms 'heat₁' and 'heat₂' are introduced for the physical properties presented in these two ways, the thought that something is hot₁ if and only if it is hot₂ can be informative. Correspondingly, consider the two thoughts that an object, presented in a given way, is hot₁ and that it is hot₂. The canonical commitments of the first thought are all fulfilled if and only if the canonical commitments of the second thought are: but this is an *a posteriori* truth (1986, p. 24).

This passage mirrors the one quoted above from Frege. The examples that Peacocke gives of canonical grounds and commitments invariably relate to the way in which the reference of the concept is identified. If the discovery that the canonical commitments have not been met obliges one to withdraw a judgment (as in the 'cube' example) this must surely be because the object of the judgment has been discovered not to instantiate the concept. A certain physical object has been discovered not to be cubic, for example, when it is discovered not to look the way it was expected to look from a different angle. The 'spectrum' of canonical commitments which Peacocke claims to be associated with a concept are therefore components which combine to constitute a Way of Referring to the reference of the concept. Peacocke's view thus implies that the individuation of a concept makes essential reference to a particular Way of Referring. This is particularly clear in the cases of 'heat₁' and 'heat₂' in the quoted passage above, where each of these concepts is held to be constitutively associated with a different Way of Referring to the same thing (one Way of Referring involves the recognition of heat through tactile sensations whereas the other involves recognising heat visually).

Another variety of Inferential Role Semantics is found in theories drawing on the ideas behind David Kaplan's (1978, 1989, 1989a, 1990) 'dthat' term. Kaplan's distinction between 'content' and 'character' is well known. To avoid confusion with other uses of the word 'content' I shall spell it 'kontent' in what follows whenever I am using it specifically in Kaplan's sense (except in quotations).¹³ The character of a term determines its kontent for a given context. The kontent of a sentence is the Russellian proposition expressed by the sentence. But it is character which determines the cognitive significance of a term. The character of a term thus corresponds to the mode of presentation associated with the term. And, since the kontent of a referring term is the item referred to, the mode of presentation once again determines the reference, albeit with context taken into account (where relevant).

Now, Kaplan introduced his term 'dthat' as a means of formalising the way in which a definite description can be used in referring directly to something. His idea was that sometimes a description could take the place of a demonstration; it is a kind of specialised 'pointing'. This allows there to be direct reference to an object using a definite description to single out the object but without the description being part of the kontent thus determined. Thus the expression 'dthat (the ϕ)' refers directly to some object which is picked out by virtue of satisfying the definite description 'the ϕ ' but the description is not part of the kontent expressed by 'dthat (the ϕ) is F'.¹⁴ The definite description thus merely 'fixes the reference' (Kripke 1980) of the term. As Kaplan (1989a, p. 581) puts it:

The word 'dthat' was intended to be a surrogate for a true demonstrative, and the description which completes it was intended to be a surrogate for the completing demonstration. On this

¹³ This is also the way Ruth Millikan (e.g. 1993) uses the word 'kontent'.

¹⁴ In his 'Afterthoughts' (1989a) Kaplan notes that his use of 'dthat' in 'Demonstratives' (1989) was ambiguous and at times varied between two possible uses of the term. One possibility was that it could be taken as a rigidifying operator such that for any definite description 'the ϕ ' the expression 'dthat (the ϕ)' would turn 'the ϕ ' into a rigid designator. The other possibility was the one described above.

interpretation 'dthat' is a syntactically complete singular term that requires no *syntactical* completion by an operand. (A 'pointing', being extralinguistic, could hardly be a part of syntax.) The description completes the *character* of the associated occurrence of 'dthat', but it makes no contribution to content. Like a whispered aside or a gesture, the description is thought of as off-the-record (i.e. off the *content* record).

This apparatus is very useful as an aid to grasping how there could be a descriptive Way of Referring involved in thoughts containing a syntactically simple directly referring term, and indeed I shall make use of an idea like this in chapter 3. What I wish to draw attention to here, however, is where Kaplan locates the Way of Thinking; it is associated with the character of the term, and since 'the description completes the character' it follows that the description determines the Way of Thinking of the reference. So, once again, we have a notion of a Way of Thinking of the same general kind as those discussed above; an account of the individuation of a given Way of Thinking thus construed makes essential reference to a particular Way of Referring.¹⁵

There are many philosophers, especially those in the neo-Fregean tradition influenced by Evans's book *The Varieties of Reference* (1982), who take it for granted that the three broad categories of Way of Referring (demonstrative, recognitional and descriptive) necessarily correspond to different concepts. They take it for granted that 'demonstrative thoughts' and 'recognition-based thoughts', for example, are different kinds of thoughts (as opposed to merely different ways of entertaining a thought) and that consequently no demonstrative thought could contain exactly the same concepts as a recognition-based thought. Any such view is opposed to Conceptual Atomism.

Kaplan (1989a, pp. 579-581) makes it clear that it was the latter that he really had in mind in 'Demonstratives'.

¹⁵ Kaplan tends not to distinguish between thought and language when he writes about character. Recanati (1993, pp. 69-72), whose view of a mode of presentation has similarities with Kaplan's (and Perry's) argues that it is necessary to distinguish two different levels, which he calls 'linguistic' and 'psychological' modes of presentation. Perry (1977, 1979) has very similar ideas to Kaplan but in his

In addition to the philosophical views just described a certain statistical theory of concepts which holds that concepts are 'prototypes' is extremely popular among cognitive scientists. On this theory concepts are individuated by their inferential roles; that is, by the inferential relations between them and other concepts. But the inferential relations are statistical and are thus weaker than the inferential relations which would be required for concepts to be individuated by definitions, for example (i.e. by inferences which are analytic).¹⁶ The idea is that for any given concept there are certain stereotypical properties which items which instantiate that concept typically, though not always, have. A bird, for example, is typically able to fly (even though not all birds can fly), so being able to fly is part of the bundle of features statistically associated with the concept BIRD. These features, put together, build up a prototypical example of a bird, and it is this prototype which constitutes the concept. There is, in fact, substantial empirical evidence that many concepts are standardly associated with certain stereotypes; thus when asked to think of a typical vehicle people tend to think of a car, when asked to think of a typical jewel they think of a diamond, and so on.¹⁷ But this does not, of course, constitute a proof that concepts are or are individuated in terms of prototypes. The prototype theory is yet another example of the kind of epistemic, inferential role theory that opposes Conceptual Atomism.

2.4 Fodor's arguments for Conceptual Atomism

One of the few explicit advocates of Conceptual Atomism is Jerry Fodor (1998), whose views have been seen as radical and counterintuitive by many philosophers and

more recent work (1997) distinguishes 'linguistic' from 'doxastic' characters. It is of course the psychological/doxastic notions which are relevant here.

¹⁶ For a highly illuminating discussion of prototypes and what is wrong with them, see Fodor 1998, chapter 5. Fodor's own arguments against equating concepts with prototypes or recognitional capacities are discussed below, in section 3.1.

¹⁷ For a review of the relevant psychological literature see Smith and Medin 1981.

cognitive scientists.¹⁸ Fodor's motivation for endorsing Conceptual Atomism relates to the well-known importance that he attaches to a view of thoughts and thought components which is compatible with the idea that thinking (construed as the making of inferences) is computation (Fodor 1975, 1987). Fodor accepts Turing's notion of computations as truth-preserving causal relations among symbols, which of course requires there to be suitable symbols (mental representations; physical items with semantic properties) upon which computations can be performed. One does not have to accept all of Fodor's associated ideas in order to accept this general point.¹⁹ But, according to Fodor, if thinking is computation then one does at least have to accept that the notion of content must not presuppose the notion of computation; that is to say, it cannot be the case that the semantic properties of a mental representation depend on the computations that it enters into, at least not unless there is an explanation of what computation is which does not presuppose notions such as 'symbol' or 'content'. Fodor feels that no such notion of computation is likely to be found; certainly Turing's notion is not of this sort, for it presupposes the notion of a symbol.

This gives Fodor a motivation for rejecting Inferential Role Semantics. For Inferential Role Semantics seems to make it impossible to tell a computational story about what inference is without the circularity caused by also telling an inferential story about the symbols upon which the computations are performed (Fodor 1998, p. 13). But although Fodor has advocated the computational story for a long time and a form of Conceptual Atomism has probably been implicit throughout his work, it is only recently that he has made this position explicit and offered arguments for it.

Fodor offers a variety of arguments directed against various different versions of Inferential Role Semantics found in different academic disciplines. Since I largely agree with Fodor's conclusions I shall not discuss all of his arguments in detail. I do find it

¹⁸ See for example Peacocke 2000, Vision 2001.

necessary, however, to discuss an argument which he directs against the individuation of some or all concepts in terms of either prototypes or perceptual recognitional capacities (Fodor 1998, chapter 5; 1998a). This argument is an important part of Fodor's case for Conceptual Atomism but unfortunately, much as I am sympathetic to its conclusion, I do not think it is valid. The argument is based on the uncontroversial principle that complex concepts are composed from simple ones. This, according to Fodor, shows that concepts cannot be prototypes because whereas concepts are compositional, prototypes are not. Thus if you have the concepts PET and FISH then you can have the complex concept PET FISH for free, but there is no obvious way that you can combine a prototypical pet (e.g. a dog) with a prototypical fish (e.g. a trout) to get a prototypical pet fish (e.g. a goldfish). Neither can you do this in reverse, which you ought to be able to do if concepts were prototypes because having the concept PET FISH requires you to have the concepts PET and FISH.

As an argument against equating concepts with prototypes I think this is valid but it does not apply with equal validity to the individuation of concepts in terms of recognitional capacities. For it seems pretty clear that if one has the ability to recognise pets and the ability to recognise fish then all one has to do to recognise a pet fish is to decide whether it is a pet then decide whether it is a fish (cf. Peacocke 2000, pp. 337-340). Fodor is aware of this, but tries to avoid the problem by assuming that a 'recognitional capacity' is never the capacity to recognise *all* instances of the type of thing referred to, but rather the ability to recognise 'good instances' of it:

Thinking centers on the notion of an instance; recognitional capacity centers on the notion of a *good* instance. Unless you are God, whether you can recognise an instance of *X* depends on whether it's a good instance of an *X*; the less good it is, the likelier you are to fail (1998a, p. 44).

¹⁹ One does not, for example, have to accept Fodor's defence of 'methodological solipsism' (1980, 1987 chapter 2), which he himself has now repudiated (1994).

If this were correct then Fodor's argument would be valid, for it is clear that 'good instances' suffer the same problems of compositionality as prototypes. It seems to me, however, that Fodor is mistaken in limiting recognitional capacities to the recognition of good instances and, moreover, this limitation is inconsistent with Fodor's own adherence to 'informational semantics' (on which see Fodor 1987, 1994, 1998 p. 12).

It is important, first of all, to distinguish being able to recognise *any* instance of *X* from *always* being able to recognise an *X*. I cannot always recognise red things; there are circumstances under which a red thing may not look red to me and I may judge it to be some other colour. But this is quite different from the claim that there are red things which I could not, under any circumstances, recognise as red. Since, as a matter of fact, I do have the ability to recognise red things I suggest that any object which does not look red to me under any circumstances is not red.

Fodor goes wrong because he tacitly assumes that, for example, the concept FISH refers to fish of all kinds even though someone might only be able to recognise certain kinds of fish. If we abstract from the effects of deference to others this cannot be right. Maybe, as a matter of fact, my ability to recognise fish is not comprehensive; I judge dolphins to be fish and I am a bit unsure about eels. I can, however, defer to experts who can correct me on this. A defender of recognitional concepts, however, will probably conclude from this that FISH is not a recognitional concept (where a recognitional concept is one an account of whose individuation makes essential reference to a particular recognitional Way of Referring). Purely recognitional concepts are not supposed to involve deference to others; to possess the putative recognitional concept *X* it is necessary to have the ability to recognise instances of *X* without assistance. In that case, surely, the reference of my concept *X* is whatever I am (under the right circumstances) disposed to recognise as an *X*. Not, of course, whatever I judge to be an instance of *X* under *any* circumstances; this would leave no room for error. But if my use of the concept RED rests on a recognitional capacity then, as I said above, any object which I never judge to be red under normal circumstances is

not red. By the same token if I have a concept FISH for which my only Way of Referring is a recognitional capacity which picks out only 'good instances' (such as trout, but not eels) then this concept does not refer to fish at all: it refers only to certain kinds of fish. All of this follows, incidentally, from Russell's Principle.

Now Fodor, as mentioned above, advocates informational semantics. Roughly speaking, this is the view that a concept refers to a certain item (object or property) because it carries *information* (in the sense of Dretske 1981) about that item. Just as the sound of a smoke alarm carries the information that the building is on fire a tokening of RED (as in 'that is red') carries the information that there is something red in the vicinity. This requires, of course, that tokenings of RED are normally caused by the presence of red things. If a given thing is picked out by a recognitional Way of Referring associated with a concept then the presence of that thing will typically cause tokenings of the concept, so informational semantics and the notion that reference can be determined by a recognitional capacity are perfectly consistent with one another.

Now, if the recognitional capacity associated with the putative recognitional concept *X* is only sensitive to good instances of *X*, and if there is no other Way of Referring associated with the concept *X*, then it must be only good instances of *X* that cause tokenings of *X*. But then, according to informational semantics, the concept *X* does not refer to instances of *X* at all, but only to certain instances of it (the 'good' ones). My tokening of the concept FISH would be caused by the presence of trout but not by the presence of eels which implies, according to informational semantics, that this particular concept FISH refers to trout but not to eels. But if the reference is determined in this way then there is no problem with compositionality; recognising an instance of PET FISH requires only recognising that it is a pet and that it is a fish. So Fodor's argument against recognitional concepts does not work.

2.5 The dynamic argument for Conceptual Atomism

Although Fodor's arguments are not conclusive a more general and conclusive argument for Conceptual Atomism can be found by taking into account the rational relations that can exist between thoughts entertained by a single thinker at different times. Now, for many concepts it is necessary to consider whether the concept in question is possessed only by one person or whether it is shared by others. If a concept is shared with other people it may then be necessary to consider whether the sharing of the concept has any effect on the principles of its individuation. For our present purposes this would introduce an unfortunate additional complexity which could cloud the issues. Consequently I shall start by restricting my discussion to concepts which are possessed only by a single individual, leaving the discussion of shared concepts until the next chapter.

This initial restriction should not be understood to imply a prior commitment to any view about the relation between thoughts and shared languages, such as whether one is possible without the other. There are, of course, a number of philosophers who have argued that genuine thinking is only possible within a linguistic community (e.g. Davidson, Kripke 1982), and that Robinson Crusoe (had he been entirely alone on a desert island from birth) could not have had genuine thoughts. Fortunately, it is not necessary to enter into this dispute because members of a linguistic community entertain many thoughts which they do not actually express, and in some cases they may have no public language words with which the thought could be expressed.

Imagine, for example, that a feral cat comes into my garden. On seeing the cat I can entertain a singular thought about it. If this thought were to be expressed it would be most natural to express it using a demonstrative. So, for example, I might think:

(3) I have never seen that feral cat in the garden before

If the thought was left unexpressed, however, then the particular token of the demonstrative 'that feral cat' would never occur as an item of a spoken language. Yet the mode of presentation would surely still be available whether or not the demonstrative was uttered. Furthermore, the cat might start to make regular appearances in the garden and I might develop the ability to recognise it every time I saw it. I could then start to think about the cat even when it was not present, satisfying Russell's Principle in virtue of the acquired recognitional capacity. I might therefore think:

(4) I wonder whether that feral cat will visit today

Equally, I might invent a name for the cat:

(5) I wonder if Fergal will visit today

Now, this name *could* be introduced into common currency so that it could be used in communication with others. But it does not seem remotely plausible that this has to occur in order for me to be able to think singular thoughts about Fergal. By the same token, it would not be plausible to suggest that an explorer who travelled alone to an unpopulated foreign land and returned having invented names for the places, flora and fauna discovered there would not have been entertaining thoughts expressible using those names prior to returning to the linguistic community. So although it could in principle be argued that membership of a linguistic community is required in some general sense in order to be able to think (though I am not advocating this position myself), it is surely not plausible that each and every thought component must correspond to a linguistic term that is in common currency; apart from anything else, it would make it mysterious how new terms get introduced in the first place.

Let us now return to an argument discussed above for the association of Ways of Thinking with Ways of Referring. This argument had the following general form: suppose I identify an object in two different ways, corresponding to two different Ways of Referring W_1 and W_2 . Then it is always possible for me not to realise that the object identified in way W_1 is the same as the object identified in way W_2 . But whenever there is an informative identity there are two different Ways of Thinking of the object, so the two different Ways of Referring must correspond to two different Ways of Thinking. Hence Ways of Thinking must be individuated in terms of Ways of Referring.

There is clearly something in this argument which is correct. Indeed, I entirely accept that when an object is referred to in two different ways *at the same time* then if it is informative for the subject that the objects referred to in those two different ways are one and the same then there are two different Ways of Thinking of the object, that is, two different concepts. Consequently I must accept that there is a sense in which a Way of Thinking can be individuated in terms of the Way of Referring with which it is associated at a given time.

What I do not accept, however, is the inference from this to the conclusion that it is constitutive of a particular Way of Thinking that it is associated with a particular Way of Referring. Clearly this conclusion does not follow deductively from the premises stated so far, for no account is taken of the possibility that Ways of Thinking could be associated with different Ways of Referring at different times.

In fact a simple application of the Transparency Principle shows that there are circumstances under which a concept possessed at one time is identical to a concept possessed at a different time but associated with a different Way of Referring. Consider a case in which a demonstrative Way of Referring is exchanged for a recognitional or descriptive Way of Referring. Returning to the example given above, when I see the cat in my garden for the very first time my perception of the cat affords me a

demonstrative Way of Referring to it. I might then make some judgment about it, for instance:

(6) That feral cat is younger than the cat that lives next door

In this particular example a complex demonstrative, 'that *feral cat*' (as opposed to a simple demonstrative 'that') is used. This may imply that the demonstrative is supplemented with an element of description; the fact that the object in question is a feral cat helps me to avoid confusion with other objects in the vicinity. It is not crucial to our current concerns, however, whether or not there is any supplementary descriptive or recognitional element; all that matters is that we are considering a Way of Referring which is predominantly demonstrative.

Now, suppose that while looking at the cat I notice that it has certain very distinctive features. I might notice these features in such a way that I am able to describe them explicitly; I might be able to say for example that the cat has two black spots on its tail. On the other hand, I may simply acquire the ability to recognise it without being able to say how, just as we are able to recognise human faces without being able to say how we do it. Let us also suppose that I live on an island on which I know that there are only three cats and that I have seen the other two cats already and I know that they do not have the distinguishing features which I currently perceive. These distinguishing features therefore allow me to distinguish the cat I am now perceiving from all other things. They afford me a Way of Referring to the cat which is independent of my perception. This Way of Referring could be descriptive, recognitional or a combination of the two. This kind of transition from a demonstrative to a recognitional or descriptive Way of Referring is, of course, extremely common in everyday life; it is almost always necessary to make such a transition in order to think about something which is no longer perceived.

After the cat has left I can continue to think about it using the new Way of Referring. I can think of it as 'that feral cat' or I can name it 'Fergal' as described above (remember that for the time being we are supposing that if I invent any names I keep them to myself). Now, if I continued to think about the cat under these circumstances I would, *ceteris paribus*, be rationally obliged to continue to predicate the same things of it.²⁰ In other words, given my earlier thought, (6), it would be irrational of me to dissent from the thought that:

(7) That feral cat is younger than the cat that lives next door

when this thought occurred after the perception of the cat had ceased. My thinking would occur in one continuous episode between seeing the cat and continuing to think of it and hold things true of it as it went out of sight. I would simply continue to think of the cat as the same one throughout. And if, once it was out of sight, I decided to adopt the nickname 'Fergal' for it then it would be equally irrational of me to dissent from the thought that:

(8) Fergal is younger than the cat that lives next door

Let us put this a little more abstractly. I started at time t_1 with a thought of the form ' a is F '. At a later time t_2 the fact that I earlier assented to ' a is F ' made it irrational for me to dissent from a certain thought of the form ' b is F '. ' a ' corresponds to the first token of 'that feral cat' and ' b ' corresponds to the second token (or the token of 'Fergal'). There are, I suggest, only two ways that this can be explained. The simplest, which I claim to be correct, is that ' a ' and ' b ' both express the same Way of Thinking (the same concept). That way, it is easy to account for the irrationality of dissenting

²⁰ The '*ceteris paribus*' is to exclude cases in which I simply forget or change my mind about whether the predicate applies. See chapter 1, above.

from (7) or (8); for on this view (7) is the same thought as (6) and one should not assent to a thought at one moment and dissent from it a moment later without a reason. And since (8) stands in the same rational relation to (6) the same point applies; there must be a single concept expressed by both 'that feral cat' in (6) and 'Fergal' in (8).²¹ In other words, someone who thinks that '*a* is *F*' at t_1 would, *ceteris paribus*, be irrational to dissent from '*a* is *F*' a moment later; so the suggestion that in fact '*a*' and '*b*' express the same mode of presentation explains things very easily. This is, in effect, a straightforward application of the Transparency Principle.

The alternative view says that the Way of Thinking of the cat changes as soon as the Way of Referring to it changes (i.e. as soon as the cat goes out of sight and a demonstrative identification is no longer possible). In our more abstract notation the suggestion is that '*a*' and '*b*' express different Ways of Thinking. But it is necessary to explain the rational relation which exists between the thought at t_1 and the thought at t_2 (why assent to the former makes it irrational to dissent from the latter). The only way that this can be explained is for there to be an identity judgment of the form '*a* = *b*' so that the thoughts at t_1 and t_2 are related by an inference of the form: '*a* is *F*; *a* = *b*; therefore *b* is *F*'.

Now, in order for the identity '*a* = *b*' to be a possible object of thought it would be necessary to entertain the different Ways of Thinking expressed by '*a*' and '*b*' at the same time. This means the cat would be thought of in two different ways at once: demonstratively and either recognitionally or descriptively. There would of course be a temporal overlap in the two Ways of Referring; since we are supposing that the recognitional capacity or descriptive information is gained from the perceptual encounter with the cat there must be a period of time during which this new Way of

²¹ One slight qualification: I am assuming that 'Fergal' and the demonstrative in (6) express the same singular mode of presentation. But it need not be the case that (6) and (8) express exactly the same overall thought, because (6) involves predicating of the object in question that it is a feral cat whereas (8) does not. This qualification would have been unnecessary if the demonstrative in (6) were not complex. For more on the nature of the thoughts expressed using complex demonstratives see chapter 6.

Referring has been acquired but the cat is also still identified demonstratively. So it could be very tempting for the theorist who holds that there is a constitutive relation between Ways of Thinking and Ways of Referring to suppose that there are indeed two different Ways of Thinking of the cat at once, and that the rational relation between the earlier and later thoughts is due to a judgment of identity between the two Ways of Thinking of the cat:

(9) That feral cat (thought about demonstratively) = that feral cat (thought about recognitionally/descriptively)

This, however, cannot be correct. It is not plausible that there must be two simultaneous Ways of Thinking of the cat in the situation described. Ways of Thinking are supposed to be psychologically real; where there are different Ways of Thinking an identity between them is supposed to be at least potentially informative. But there is no possibility of an informative identity between two different Ways of Thinking of the cat in the circumstances described (which are simply that I see a cat and notice its distinguishing features). If there were two different Ways of Thinking, 'a' and 'b', then it ought at least to be possible to make sense of a suggestion that 'a is F but b is not F'. But it is not possible to make sense of such a suggestion in the case in question. There is no way to make sense of a thought such as:

(10) That feral cat (thought about demonstratively) is younger than the cat that lives next door, but that feral cat (thought about recognitionally/descriptively) is not younger than the cat that lives next door

It is not merely that such a thought would be obviously contradictory; rather, in the circumstances described it does not seem possible to entertain such a thought at all.

It is important when thinking about this matter to make a distinction between a judgment of identity and a judgment that an object has a certain property. Consider the distinction between the following two judgments:

(11) a is the ϕ

(12) $a = b$ (where b is identified descriptively as 'the ϕ ')

A similar pair, using Kaplan's terminology, would be:

(13) that is the ϕ

(14) that = dthat (the ϕ)

The first judgment in each pair ascribes a property to an object; it says that the object is uniquely ϕ . The second judgment, by contrast, is an identity statement; an object thought about in one way is identified with an object thought about in another way. Two possible analyses of (11), for example, would be:

(15) $\exists x [a = x \ \& \ \phi x \ \& \ \forall y (\phi y \rightarrow y = x)]$

(16) $\phi a \ \& \ \forall y (\phi y \rightarrow y = a)$

I suspect that (16) is a better analysis since intuitively (11) does not make any assertion about existence, but in any case both versions clearly differ from an identity such as (12). A similar distinction could be made between (13) and (14). It is essential to be able to make distinctions of this kind, for otherwise there would be no such thing as ascribing properties to objects. There would only be judgments of identity between differently identified objects, a possibility whose implausibility I take for granted.

In the example of the cat it is clear that in acquiring a recognitional capacity or a descriptive Way of Referring I make a judgment ascribing properties to the cat rather than identifying a demonstratively identified cat with a recognitionally or descriptively identified cat. For example, I might first see the cat, then judge that:

(17) That feral cat has two black spots on its tail

This is an entirely different judgment from, for example:

(18) That feral cat (demonstrative) = that feral cat (the one with two black spots on its tail)

Whereas (18) involves two different Ways of Thinking of the cat, (17) is a subject-predicate judgment involving only one Way of Thinking. The property that is ascribed in (17) can, however, be used to identify the cat once it has disappeared from sight. The same principle applies to the acquisition of a recognitional capacity except that there is no articulated ascription of a property (at most there could be an ascription of the property of 'looking like *that*'). Hence, strictly speaking, the acquisition of a recognitional capacity need not involve a judgment at all, but this only strengthens my point.

To sum up, we started from the observation that assenting to a thought of the form '*a* is *F*' can make it irrational to dissent from a thought of the form '*b* is *F*' at a later time even though '*a*' and '*b*' are associated with different Ways of Referring. In our example '*a*' is associated with a demonstrative Way of Referring and '*b*' is associated with either a recognitional or a descriptive Way of Referring. We then noted that if the Way of Referring associated with '*b*' was acquired from the perceptual episode which provided the demonstrative Way of Referring associated with '*a*' then the rational relation between '*a*' and '*b*' is not explained by an identity judgment '*a* = *b*'. It follows

that 'a' and 'b' must express the same concept. And since 'a' and 'b' are associated with different Ways of Referring it cannot be the case that concepts are individuated in terms of Ways of Referring. The latter can change while the former remains the same. Consequently an account of the individuation of a given concept makes no essential reference to any particular Way of Referring; Conceptual Atomism is correct.

Although I have used an example where a demonstrative Way of Referring is exchanged for a recognitional or descriptive Way of Referring the same sort of case can be made whenever there is a transition from one Way of Referring to another. There could, for example, be changes in the descriptive or recognitional information used to identify the reference which would allow the kind of psychological continuity through time which requires there to be a continuous Way of Thinking. I have discussed only terms which refer to objects but nothing I have said is specific to any one kind of term. Changes in the Way of Referring occur more rarely with terms which refer to properties but they can nevertheless occur for some concepts and a similar argument can be constructed. There may be some concepts for which there is only one possible Way of Referring to the reference in which case no comparable argument about the dynamics of such concepts can be constructed. If Conceptual Atomism offers a satisfactory account for all concepts, however, whereas its rivals fail for all but certain special cases then it seems reasonable to accept Conceptual Atomism as an account of all concepts.

2.6 *Changing one's mind*

The persistence of the identities of various things through time is a fundamental metaphysical feature of the world.²² It is required in order for there to be such a thing

²² In saying this, I do not intend to take any position on the endurance/perdurance debate in temporal metaphysics (Lewis 1986); both notions are consistent with the present point.

as change because in order for it to make sense to say that something changes, something must persist.²³ This is reflected in the fact that in order for one to change one's mind about the nature of something, rather than merely changing one's beliefs, there must be a mode of presentation of it which persists while what is predicated of it changes.

This is, in effect, another way to express the key idea behind the dynamic argument for Conceptual Atomism. For the majority of concepts it is possible to change one's mind about almost any property which one predicates of the concept without changing the concept.²⁴ This includes the properties which one uses at a given time to identify the reference of the concept, so long as they can be replaced by some other Way of Referring. This applies to recognitional capacities as well as to descriptive identifications; if, for example, someone identifies me in virtue of a capacity to recognise my face they will have to update this recognitional capacity as I get older and my appearance changes. I may end up looking quite different to the way I looked when I was younger and the recognitional capacity will make use of quite different features of my appearance. Yet there is still the same mode of presentation of me throughout; so long as the updating process is reasonably continuous, allowing a gradual transition of the recognitional capacity (so that there are no intervals between encounters with me which are long enough that I might not be recognised), I will simply be thought of as the same person, looking gradually older.

The relation between changing one's mind and the retention of a concept was drawn attention to in the context of indexical thoughts by David Kaplan (1989, pp. 537-8):

I want briefly to raise the problem of *cognitive dynamics*. Suppose that yesterday you said, and believed it, 'It is a nice day today'. What does it mean to say, today, that you have

²³ This point is made by many philosophers including Kant in the First Analogy of the *Critique of Pure Reason* (1933) where it is used in an argument for the existence of persisting 'substance'.

retained *that* belief? It seems unsatisfactory to just believe the same content under any old character - where is the *retention*?... What does it mean to say of an individual who at one time sincerely asserted a sentence containing indexicals that at some later time he has (or has not) *changed his mind* with respect to his assertion? (p. 537n).

The distinction that Kaplan has in mind is between sequences of thoughts such as the following:

- (19) At t_1 : Cicero was a great orator
- (20) At t_2 : Cicero was not a great orator

- (21) At t_1 : Cicero was a great orator
- (22) At t_2 : Tully was not a great orator

Let us assume that 'Cicero' and 'Tully' both refer to the same person but that the person who thinks the above thoughts does not know this. In that case there is a sense in which the sequence (19)-(20) can represent a change of mind but (21)-(22) cannot. In both cases the same content is entertained at t_2 but there is clearly an important difference at the level of thought. Changing one's mind cannot just mean changing from saying, of something, that it is F to saying, of that same thing, that it is not F (purely at the level of content) because that is what happens in both of the cases just described but only one of them involves changing one's mind. Kaplan assumes that differences like those between (19)-(20) and (21)-(22) must be explained at the level of character, but he has certain difficulties in making his account work.²⁵

²⁴ The exceptions include concepts like BACHELOR, where it does not appear possible to change one's mind about whether bachelors are unmarried men. As I argue below (section 3.3), however, the existence of such concepts does nothing to refute conceptual atomism.

²⁵ Kaplan discusses cognitive dynamics with regard to indexicals but in fact examples like the Cicero/Tully one present Kaplan with an even bigger problem than indexicals. This is because the character of a proper name is just a function assigning it a reference, which means that 'Cicero' and 'Tully' have the same character. Kaplan was aware of this problem even at the time: 'I claimed that a

A different account of concept individuation solves the problem, however; in (19)-(20) there is a change of mind because at t_1 something is predicated of the concept CICERO and at t_2 its negation is predicated also of CICERO. In (21)-(22), by contrast, the concepts to which the predicates apply change between t_1 and t_2 so the former thought is not understood as having any bearing on the latter.

Now, suppose that at t_1 while the cat is in sight in the garden I think:

(23) That feral cat (identified demonstratively) is in the garden

Then at t_2 , as soon as the cat leaves the garden and simultaneously goes out of site, I think:

(24) That feral cat (identified recognitionally/descriptively) is not in the garden

The suggestion made above implies that the sequence (23)-(24) is like the Cicero-Cicero sequence (19)-(20) rather than the Cicero-Tully sequence (21)-(22).²⁶ The second thought, (24), has the kind of rational bearing on the earlier thought, (23), that makes it possible to describe the sequence as a change of mind rather than a sequence

competent speaker knows the character of words. This suggests (even if it does not imply) that if two proper names have the same character, the competent speaker knows that. But he doesn't.' (1989 pp. 562-3).

²⁶ Strictly speaking, (23) and (24) should be understood to imply that the cat is or is not *currently* in the garden. Since the two thoughts occur at different times they therefore do not contradict one another. But there is still a relevant sense in which this counts as a change of mind. The later thought still has the relevant bearing on the earlier thought; I am aware that what I think now is a *change* from what I thought about that (same) feral cat earlier on. Consider, as a further example, the difference between the sequences 'Cicero is alive; Cicero is dead' and 'Tully is alive; Cicero is dead'. In the former sequence the new piece of information (that Cicero is dead) has a bearing on the preceding thought (that Cicero is alive) even though, if the thoughts occur at different times, the later thought does not strictly contradict the earlier one. There is no such relation between the thoughts in the latter sequence. I could, of course, have chosen a predicate for (23) and (24) for which there was a strict contradiction; the current one, however, helps to illustrate the point I am making about the transition from a demonstrative to a recognitional Way of Referring.

of unrelated thoughts.²⁷ Consequently a concept is retained from (23) to (24). But, to reiterate the argument, only Conceptual Atomism can reconcile this retention with the fact that the concepts in (23) and (24) involve different Ways of Referring.

2.7 Reason and Perception

There are surprisingly few explicit arguments in favour of Inferential Role Semantics apart from the standard Fregean one discussed above.²⁸ There is, however, an argument which has been expounded in recent years by Christopher Peacocke (1992, 2000) to which it is necessary to respond in detail because it has in common with my own argument the fact that it appeals to the need to individuate concepts in such a way as to make rational sense of the thinker.

Peacocke describes his view as a version of Fregean Inferential Role Semantics (IRS). He characterises two key features of such theories (2000, p. 332):

- (I) We need in psychological explanation and in epistemology a notion of a way of thinking individuated by the Fregean Criterion of informativeness.
- (II) The notion of a way of thinking motivated by (I) is either individuated by, or at least is constrained in various constitutive ways, by its conceptual or inferential role.

Principle (I) is what makes a theory Fregean. Principle (II) is what makes it a version of Inferential Role Semantics. Peacocke (2000, p. 332) explains the relation between these two principles and the rationality of the thinker as follows:

²⁷ Michael Luntley (1998, 1999 chapter 12) makes some related points about indexicals. See below, chapter 5.

What all Fregean IRS theories, of whatever stripe, agree upon is the central place of reasons and rationality in the individuation of concepts. Reasons for making judgments are central in any Fregean theory, since the informativeness criterion appeals to what can be reasonably judged in given circumstances. The distinctive claim of the Fregean IRS theorist is then that we can give a philosophical explanation of how reasons contribute to the individuation of concepts by appealing to inferential or conceptual role.

Now, this assumes that there is something more to the ‘inferential role’ of a term than is implied by principle (I). To see what is implied by principle (I), for example, consider the fact that an alternative way of capturing the Fregean informativeness criterion would be to appeal to the notion of formally valid inference.²⁹ To say that ‘ $a = b$ ’ is informative while ‘ $a = a$ ’ is not could, for example, be seen as equivalent to saying that (25) is a valid inference while (26) is not:³⁰

(25) $\vdash a = a$

(26) $\vdash a = b$

There is, of course, a notion of ‘valid inference’ according to which an inference is valid whenever it is the case that if the assumptions are true the conclusions are true. This would make (26) a valid inference. But another notion of validity is relevant here; an inference is valid, and should be drawn by any rational person, if it has a valid form. (25) is valid in this sense because whatever ‘ a ’ refers to an inference of that form is

²⁸ Fodor (1998, 2000) notes at some length how surprisingly few explicit arguments there are in favour of Inferential Role Semantics even though some version of the position is often taken for granted in both philosophy and cognitive science.

²⁹ This way of individuating concepts (sense) is described by John Campbell 1987, 1994, pp. 73-82. Timothy Williamson (1997) registers an objection, but see Campbell’s (1997a, pp. 664-669) response.

³⁰ I use the symbol ‘ \vdash ’ as the ‘assertion sign’ (Lemmon 1987, pp. 11-12): assumptions are listed on the left of the sign and the conclusion which is held to follow from those assumptions is stated on the right. (25) and (26) are therefore being claimed to follow from no assumptions (incorrectly in the case of (26)).

valid. We could substitute any other term for 'a' in (25) (in both positions) and obtain a valid inference. This cannot be said for (26).

It will perhaps be clearer that this point concerns inference in general and not just judgments of identity if we consider another example:

(27) a is F , a is $G \vdash a$ is both F and G

(28) a is F , b is $G \vdash a$ is both F and G

Once again, if ' $a = b$ ' is true then in both cases true assumptions entail a true conclusion. But only (27) is formally valid; someone who does not know that $a = b$ can rationally make the transition represented by (27) but not the transition represented by (28). (27) is formally valid because all of the occurrences of ' a ' express the same mode of presentation whereas (28) is not formally valid because ' b ' expresses a different one.³¹ This is what the individuation of modes of presentation has to capture. This is one sense, then, in which any Fregean should agree that modes of presentation must be individuated in such a way as to capture their 'inferential roles'. Let us call an inferential role, construed as that aspect of the role of a term which is required for inferences containing it to be formally valid, a *formal inferential role*. This is a very insubstantial notion of a 'role', for terms of the same category (such as names) do not differ in their formal inferential role (if all of the ' a 's are substituted with ' b 's, for example, the validity of the inference is not changed). The notion is therefore far too coarse-grained to be used to individuate concepts.

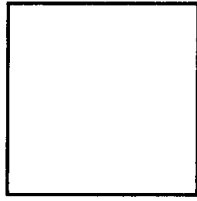
³¹ I am not assuming that modes of presentation can be individuated by symbol type; I do not deny that different tokens of the same symbol ' a ' could be used in expressing different modes of presentation. I shall, however, stipulate that my terminology avoids this. If an inference can be written down with the same form as (27) without being valid then this is because different tokens of the same symbol type are in fact being used to express different modes of presentation. But if we adopt a terminology which makes transparent what is going on at the level of thought then, by stipulation, this will not occur. See, once again, Williamson 1997 and Campbell 1997a, and see also Luntley 1999, pp. 303-312.

All of this, as Peacocke seems to tacitly acknowledge in the quotation above, is implied by principle (I). Principle (II), however, says that modes of presentation as required by (I) can be individuated (or have their individuation constrained) by their inferential role. If this is not trivial then there must be more to the inferential role of a term than its formal inferential role. Peacocke thinks there is; moreover, he thinks that terms which are associated with different Ways of Referring have different inferential roles and it is this which makes them different concepts.

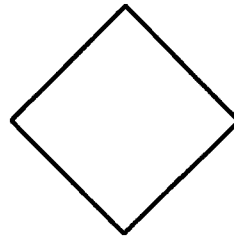
Peacocke does not apply this reasoning to all differences in Ways of Referring. He shows signs, for example, of accepting the idea that there could be changes in descriptive Ways of Referring without changes in the concept:

Recognition of the level of concepts in the present sense is also needed for a proper description of cases in which there is a change of prototype or of relatively central beliefs associated with a concept. Indeed, as elsewhere, describing a case as one of change, rather than as one of replacement, is correct only if there is something that persists through the change, and it is the concept in my sense that so persists (1992, p. 3).

Peacocke's main focus, however, is on recognitional and demonstrative Ways of Referring and the concepts which he takes to be constitutively related to these. To take one of Peacocke's own favourite examples, he argues that it is possible for someone to possess different concepts SQUARE and REGULAR DIAMOND which refer to the same shape property. Moreover, he argues that when presented with an object in the orientation of figure 1a (below) it may be rationally required for the subject to judge that 'that is a square' rather than 'that is a regular diamond', and vice versa for figure 1b. It is, Peacocke claims, constitutive of these concepts that someone who possesses them makes these judgments as described and not, for example, the other way around.



1a



1b

Strictly speaking, it is not the orientation of the shapes that matters but rather the symmetries to which the subject is sensitive. Peacocke (1992, pp. 74-7) argues that perceiving something as square requires an awareness of the symmetries around lines parallel to its sides whereas perceiving something as a regular diamond requires an awareness of its symmetries relative to its diagonals. Consequently it is in fact possible for most of us to perceive figure 1b as square and figure 1a as a regular diamond; the difference in the two orientations simply makes certain symmetries more salient. There could, however, be a subject who only noticed the square-symmetries for shapes in the orientation of figure 1a and only noticed the regular diamond-symmetries for shapes in the orientation of figure 1b. If a subject of this kind was presented with figure 1a it would, according to Peacocke (2000), be rationally obligatory for that subject to judge that ‘that is a square’ rather than ‘that is a regular diamond’, and vice versa for figure 1b.

It is important to realise that SQUARE and REGULAR DIAMOND are what Peacocke calls ‘observational concepts’. They are not the geometrical concepts usually expressed using the same names; for, of course, anyone with the geometrical concept of a square might check on whether something was square by making measurements rather than by simply looking, and if they did they would find that both 1a and 1b are (approximately) square. Rather, these are supposed to be concepts which pick out a certain shape just because of the way it appears to the subject. The extensions of the observational and geometrical terms need not be the same. Consequently, for example, something might

be correctly judged to be square in the observational sense which is not precisely square in the geometrical sense. Roughly speaking, something is square in the observational sense if and only if it appears square to a normal observer under normal conditions. The same applies, *mutatis mutandis*, to regular diamonds; and Peacocke is presumably considering a subject whose perceptual sensitivities are such that they pick out the very same shape property with SQUARE and REGULAR DIAMOND, the difference in the concept judged depending only on the different ways in which the shape is identified (the difference in the symmetries to which the subject is sensitive).³²

Since, Peacocke argues, a subject with the dispositions described above is rationally obliged to judge that 'that is a square' rather than 'that is a regular diamond' on being presented with 1a, and since concepts are to be individuated so as to respect such rational considerations, it must be the case that SQUARE and REGULAR DIAMOND are different concepts; each of them being constitutively associated with a different recognitional Way of Referring. Being confronted with 1a gives the subject a reason to judge that 'that is a square' and being confronted with 1b gives the subject a reason to judge that 'that is a regular diamond' but not vice versa. The individuation of concepts must respect the reasons the subject has for accepting a content containing a given concept, so the rational relation just described constrains the individuation of the concepts.

Individuating concepts in this way does not merely add an extra fineness of grain to the principles of rational individuation that were applied in the diachronic cases described above. On the contrary, despite appealing to the same grounds of rationalising individuation, Peacocke's theory is in direct conflict with the diachronic individuation of concepts for which I have been arguing. It says that certain concepts

³² SQUARE and REGULAR DIAMOND do not, however, refer to a relation between the shape and the observer; i.e. it is *not* the case that something is square only if it is in the correct orientation with respect to the observer. This would be an example of what, in chapter 6, I call an 'egocentric predicate'. Rather, the reference of the concepts currently under discussion is simply a mind-independent shape, which can be picked out in two different ways (i.e. using two different perceptual dispositions).

must be different where my theory says that they must be the same. In my opinion, Peacocke's individuation of concepts offers us no explanation of many diachronic cases of the kind discussed above. By the same token, however, in order for my theory to be acceptable I must show where Peacocke goes wrong.

Peacocke's argument requires there to be a sense in which SQUARE and REGULAR DIAMOND differ in their inferential roles, where there is more to an inferential role than a merely formal inferential role. SQUARE and REGULAR DIAMOND can be seen to be different concepts from the fact that (29) is a valid inference whereas (30) is not:

(29) f is square, g is square \vdash both f and g are square

(30) f is square, g is a regular diamond \vdash both f and g are square

But Peacocke thinks that the difference between SQUARE and REGULAR DIAMOND has to do with the difference in their 'inferential roles', the supposed difference between whether a rational subject should apply one concept or the other in certain perceptual situations, as described above. If it can be shown, however, that there is nothing more to the inferential role of a term than its formal inferential role then Peacocke's attempt to explain the latter in terms of the former clearly fails.

I shall first show that Peacocke's example of SQUARE and REGULAR DIAMOND fails to prove his point then I shall give some more general arguments against his view. In order to be able to compare the concepts SQUARE and REGULAR DIAMOND on equal terms it is necessary to assume that REGULAR DIAMOND is a simple, unstructured concept (for which there is, as far as I know, no word in English) rather than a complex concept made up of REGULAR and DIAMOND. But because it is very natural when thinking about this example to have in mind the complex concept rather than the novel simple one, one's intuitions get rather distorted. Someone who possesses the complex concept REGULAR DIAMOND must possess the concept DIAMOND, and in order to possess the concept DIAMOND one must possess a way of distinguishing diamonds

from all other things. The concept DIAMOND is not, of course, co-referential with the concept SQUARE and in order to judge that something is a diamond it may well be necessary to be sensitive to the symmetries made salient by 1b rather than 1a. If the simple concept which we are invited to consider is called 'REGULAR DIAMOND' it is therefore likely to retain this association, which naturally pushes one's intuitions in the direction in which Peacocke wishes them to go. Let us therefore refrain from calling this the concept REGULAR DIAMOND and call it REG instead.

Now, suppose that a subject has only the concept SQUARE and not the concept REG and suppose that this subject only ever notices that something is square when it is in the orientation of 1a. Someone could take a square object in the orientation of 1a and, before the subject's eyes, rotate it through forty five degrees, saying 'look, squares can look like *this*'. The subject might, at this point, become aware of the diagonal symmetries of the shape for the first time. But this would not be a discovery that squares and regs are the same thing; rather, it would be the acquisition of a new way of identifying squares, for the rotated shape was kept track of throughout and had already been recognised as a square. So now the reference of a single concept, SQUARE, could be identified using either of the two symmetries. It would even be possible for the subject to subsequently forget about the symmetries normally associated with squares while still retaining the same concept. None of this would be possible, however, if the concept SQUARE was constitutively tied to a particular Way of Referring, as Peacocke suggests. This argument is, of course, just a particular application of the more general dynamic argument for Conceptual Atomism discussed above in section 2.5.

Just in case the intuition that the concept REG is acquired when the shape is rotated still seems compelling, consider how crazy the same reasoning would seem when applied to concepts more generally. Consider, for example, what happens when Jacques Chirac, whom I had previously only been able to recognise from the front, turns around before my eyes. I now acquire the ability to recognise Jacques Chirac by looking at the back of his head, as well as from the front. Does this mean that I acquire

a new concept JACQUES CHIRAC (SEEN FROM BEHIND) to add to my pre-existing concept JACQUES CHIRAC (SEEN FROM THE FRONT)? Of course not. SQUARE and REG should not be treated any differently.

It should be possible to construct a similar argument for any pair of co-referential concepts. If this is correct then Peacocke's argument is refuted. We can, however, consider what is wrong with views of this kind at a deeper, more general level. The general structure of the argument was as follows: firstly, it was observed that the individuation of concepts must respect the reasons that someone might have for accepting a content containing a given concept. This I accept. Secondly, however, it was claimed that particular types of perceptual experiences provide reasons for accepting certain contents. It was inferred from this that there are certain concepts whose individuation is constitutively related to particular perceptual sensitivities or recognitional capacities.

The problem with this line of argument is that it does not work if the content of experience is itself conceptual, yet unless experience is conceptual there can be no rational relation between experience and belief. If experience is conceptual then the claim that a perception as of an *F* provides a reason to believe that the perceived item is an *F* cannot be used to individuate concepts without circularity. So Peacocke's argument only makes sense on the assumption that the putative reasons are given by something extra-conceptual; presumably the supposed *non-conceptual content* of perception.³³

A number of philosophers have attacked the idea that anything extra-conceptual could stand in a rational relation to belief. Donald Davidson, for example, claims that

³³ It is only 'observational concepts' that should be considered here. Suppose that I cannot recognise dodecagons except by counting their twelve sides. Then I could have a conceptual representation of something as, for example, a shape or a polygon without immediately recognising it as a dodecagon; hence there would be no need to posit non-conceptual content as a reason for accepting that a perceived shape is a dodecagon. My judgment (based on counting) that the shape has twelve sides might of course form part of an inference which leads to my judging it to be a dodecagon. But Peacocke's idea is presumably that observational concepts are not like this; and that one is able to recognise squares, regular diamonds, red things etc. immediately and non-inferentially.

‘nothing can count as a reason for holding a belief except another belief’ (1990, p. 123). He argues as follows:

Suppose we say that sensations themselves, verbalised or not, justify certain beliefs that go beyond what is given in sensation. So, under certain conditions, having the sensation of seeing a green light flashing may justify the belief that a green light is flashing. The problem is to see how the sensation justifies the belief. Of course, if someone has the sensation of seeing a green light flashing, it is likely, under certain circumstances, that a green light is flashing. *We* can say this, since we know of his sensation, but *he* can’t say it, since we are supposing he is justified without having to depend on believing he has the sensation. Suppose he believed he didn’t have the sensation. Would the sensation still justify him in the belief in an objective green light? (1990, pp. 124-5).

Davidson presumably construes ‘sensation’ as something extra-conceptual. One cannot, of course, give a reason for believing that one has certain sensations; one just *does* believe it.³⁴ But although he does not think that sensations justify beliefs Davidson does not imagine that beliefs are adopted independently of the way things are:

The relation between a sensation and a belief cannot be logical, since sensations are not beliefs or other propositional attitudes. What then is the relation? The answer is, I think, obvious: the relation is causal. Sensations cause some beliefs and in *this* sense are the basis or ground of those beliefs. But a causal explanation of a belief does not show how or why the belief is justified (1990, p. 125).

³⁴ Peacocke (1986, p. 111), in responding to the remark by Davidson quoted above, casts doubt on whether it is intelligible for someone to have an experience of a certain subjective character yet believe that there is no such experience. Even if this is correct it does little to deflate Davidson’s overall point. But if it is correct it also lends support to the claim that ultimately some beliefs must be accepted without justification; for in order for an experience to justify my belief that I am having a certain experience there would have to be a gap between whatever it is that is supposed to justify the belief and the belief itself. That gap seems to be lacking in the case of subjective experience.

The first sentence captures a crucial point, that there is no logical relation between sensation and belief. This is clear from the fact that sensation, construed as extra-conceptual, would always underdetermine belief. As Hume (1978) showed, there is no deductive inference from any finite set of observations to a generalisation. Attempts to prove this conclusion wrong have been unconvincing. Attempts have been made, for example, to show that a finite number of observations can lend probabilistic support to a theory. But they cannot; the probability is always zero (see Miller 1994). The same principle applies to perception. In a section of *The Logic of Scientific Discovery* (1959) in which he explicitly attacks the idea that experiences justify beliefs, describing this idea as a form of psychologism, Karl Popper argues that all empirical statements involve universals which go beyond what can be observed (he describes this as ‘the transcendence inherent in any description’ (1959, p. 94)):

Every description uses *universal* names (or symbols, or ideas); every statement has the character of a theory, of a hypothesis. The statement ‘Here is a glass of water’ cannot be verified by any observational experience. The reason is that the *universals* which appear in it cannot be correlated with any specific sense-experience. (An ‘immediate experience’ is *only once* ‘immediately given’; it is unique.) By the word ‘glass’, for example, we denote physical bodies which exhibit a certain *law-like behaviour*, and the same holds for the word ‘water’ (1959, pp. 94-5).

The same could be said even about terms like ‘green’, as can be seen from the fact that a perception which supposedly gave a reason to believe that something is green would give just the same reason for thinking that it was ‘grue’ (Goodman, 1965).

McDowell (1994) has recently argued strongly against the existence of non-conceptual content and the associated idea that anything extra-conceptual could serve

as a reason for a belief.³⁵ McDowell's arguments are well-known and I shall not rehearse them in detail here, but they depend on similar principles to Davidson's; nothing extra-conceptual could possibly be the right kind of thing to impinge upon rationality so nothing extra-conceptual could count as a reason for a belief. McDowell rejects Davidson's idea that sensation is merely the cause of beliefs, however, in favour of a view according to which experience is inherently conceptual; it is a form of 'openness to reality' (1994, p. 26) in which beliefs are justified, albeit defeasibly, because conceptual experience is the right kind of thing to impinge upon rationality.³⁶ In my opinion these arguments serve as a strong refutation of Peacocke's argument for Inferential Role Semantics.

2.8 Summary

In this chapter I have argued for Conceptual Atomism on the grounds that there are situations in which a Way of Referring can be exchanged for a different one over time

³⁵ McDowell also argues explicitly against Peacocke's (1992) particular version of the claim that something extra-conceptual could be a reason for a belief. See McDowell 1994, pp. 162-70. Bill Brewer (1999, chapter 5) defends and elaborates upon McDowell's arguments and also deals specifically with Peacocke's account.

³⁶ Although I agree with much of McDowell's account I do not, in fact, agree with the claim that experience gives absolute justification to beliefs. If one accepts Popper's (1959, 1972, 1989; see also Miller 1994) overall approach to the acquisition of true beliefs then one sees that there is room for a position according to which beliefs are accepted without justification and rationality consists in avoiding inconsistency among one's beliefs and in subjecting beliefs to the most stringent testing possible. One acquires a body of true belief because one is disposed to accept one's perceptions as veridical, not because one has any reason to do so. Just as it is an inherent quality of an itch that it brings with it a compulsion to scratch, so it is an inherent quality of perceptual experience that it brings with it the compulsion to believe, but a compulsion is not the same thing as a reason. This view of the role of perception is consistent with the popular view in cognitive science that perceptual systems generate hypotheses. To accept one's perceptions as veridical (when one is 'open to reality') would be to treat all perceptual hypotheses as true unless one adopted a belief which was inconsistent with them. See Gregory 1970, 1981 for some of the original development of the idea and for an up-to-date overview of the evidence from cognitive science see Papathomas 1999. Daniel C. Dennett (1991), who occasionally hints at a Popper influence (e.g. Dennett 1975), also has a view of perception which could be construed as compatible with this. Brewer (1999) gives an interesting argument for the claim that perceptions provide reasons for beliefs; I do not believe that this is valid but unfortunately a discussion of this would involve too great a departure from the arguments of this thesis.

without a change of concept. The concept must be retained because there are rational relations between the thoughts entertained at different times which can only be explained by a retained concept. Consequently an account of what it is to possess a given concept makes no essential reference to any particular Way of Referring.

I have also refuted Peacocke's version of Inferential Role Semantics, according to which certain concepts are individuated by the rational relations in which they stand to certain kinds of perceptions. I argued against Peacocke in two ways. Firstly, I argued that his claim cannot be correct for the examples he uses (such as SQUARE and REGULAR DIAMOND), because the Ways of Referring associated with these concepts could in fact be exchanged over time. Hence their relations to perceptions of a particular type are contingent rather than constitutive. It does not seem likely that there could be any examples which would not be subject to this objection. Secondly, Peacocke's argument could only work on the assumption that perception contains non-conceptual content; an assumption which, following McDowell and others, I reject.

CHAPTER 3

Concepts II: An Atomistic Theory

The purpose of the last chapter was largely negative; its main aim was to show that a common set of views about concepts which can be grouped together under the heading of Inferential Role Semantics should be rejected because they cannot give a satisfactory account of cognitive dynamics. But rejecting a view is only ever fully plausible if it can be shown that there is an acceptable alternative. In this chapter I shall outline a version of Conceptual Atomism which has no difficulties with cognitive dynamics. According to this theory concepts are individuated in terms of *referential episodes*, episodes of ongoing reference to a particular thing during which the Way of Referring may change in various ways provided there is the diachronic transparency required by the Transparency Principle. A concept is created when a referential episode begins and ceases to exist when it ends and the difference between different concepts consists in nothing more than the difference in numerical identity between different referential episodes. Before describing this theory in more detail, however, I shall briefly discuss the views of three philosophers - Gareth Evans, John McDowell and some recent views of David Kaplan - in order to show how, despite expressing their theories in different terms, a number of philosophers have been gradually moving towards positions which are, or can be made compatible with, versions of Conceptual Atomism. This will put Conceptual Atomism in a broader context and will also help with the subsequent exposition.

3.1 *A convergence of theories*

The interpretation of Gareth Evans's (1982) theory is debatable. Evans can very easily be interpreted as individuating Ways of Thinking in terms of Ways of Referring and I suspect that many readers have understood him in this way. This may very well be a correct interpretation. There is, however, some significant ambiguity in what Evans says. I shall quote a few passages to illustrate this ambiguity, but ultimately I shall be more concerned to describe a position which Evans could have held, given many of his other views, rather than to debate whether or not he actually did hold it.

Here is Evans's official explanation of what he means by a 'way of thinking':

I suggest that the desired notion can be explained in terms of the notion of an account of what makes it the case that a subject's thought is a thought about the object in question. Imagine such an account written out. 'S is thinking about the object *a* in virtue of the fact that ...S...': what follows 'that' is an account in which references to the subject and the object thought about appear, possibly at several places. Now I suggest that another subject, *S'*, can be said to be thinking about the object *a* in the same way if and only if we get a true statement when we replace reference to *S* with reference to *S'* throughout the account provided for *S*, deriving 'S' is thinking about *a* in virtue of the fact that ...S'...' (1982, pp. 20-1).

Since the account of what it is to think in a given way involves the object thought about as well as the subject this may be a good way of capturing the Fregean idea of a mode of presentation as a semantic notion. But in itself it tells us very little about what Evans thought a 'way of thinking' was, because it leaves open just what kind of account should be given of what makes it the case that the subject's thought is about the object in question. It is, however, very tempting to read Evans's subsequent lengthy discussion of the different varieties of reference as a discussion of precisely

what forms such an account might take. This interpretation is especially tempting in the light of remarks such as this one:

To give an account of how a thought concerns an object is to explain how the subject knows which object is in question (1981, p. 289).

Thus it is very natural to interpret Evans as holding, for example, that when an object is identified demonstratively a different account must be given of what makes it the case that the thought concerns the object than when the object is identified by description or by a recognitional capacity. Hence the Ways of Thinking are different in these different cases (and in any others in which there are different Ways of Referring).

It is a little difficult, however, to reconcile this understanding of Evans with passages such as the following:

It is important to understand that the ability to keep track of an object must allow both subject and object to move during the period of observation. Note also that it would be quite arbitrary to deny that the same ability can be exercised in cases in which the object disappears momentarily behind an obstacle (1982, p. 176).

This passage suggests that Evans believed that the Way of Thinking of an object can be regarded as the same even if the object goes temporarily out of sight. This implies that an object can be identified purely demonstratively for a period of time, then recognitionally, then demonstratively again, without a change in the Way of Thinking. Consider also:

There is indeed no general guarantee that a belief will persist. Suppose I perceive an object and judge 'This is *F*', and then lose track of it. There is no guarantee that there will be available to me a past-tense demonstrative Idea..., enabling me to judge 'That was *F*'. (If

such an Idea *is* available to me, it will need to draw on conceptual material not present in the original demonstrative identification.) (1982, p. 236).

This passage states that retaining a Way of Thinking is not always possible if the object is lost track of (making a demonstrative Way of Referring unavailable). But the last bracketed sentence implies that such retention is possible if the subject has acquired a recognitional or descriptive Way of Referring during the perceptual encounter with the object, which again seems to conflict with the individuation of Ways of Thinking in terms of Ways of Referring.¹

Further support for a more 'atomistic' reading of Evans comes from his endorsement of McDowell's (1977) 'austere' construal of the way in which the senses of expressions can be given by a theory of meaning. Evans (1981) argues (following both McDowell (1977) and to some extent Dummett (1981)) that a semantic theory for a language should be based on a theory of reference. This is a theory which assigns a reference to each meaningful expression of the language in a way which reflects the compositional structure of the language.² Now there can, of course, be many different equally correct theories of reference for a given language. Two such theories could, for example, be identical in all respects except for the fact that they offered the following two different clauses in assigning a reference to the word 'Hesperus' (it is assumed that 'Hesperus' and 'Phosphorus', as they appear on the right hand side, both refer to the same object):

¹ In quoting these passages and giving them the interpretation I have described, however, I do not mean to deny that there are sections in which Evans says things which seem quite incompatible with this reading e.g. 'If this is true in the case where the Fregean sense of a singular term is purely descriptive...' (1982, p. 61). Whatever Evans may have intended, however, such passages can often be re-read as concerning Ways of Thinking that are only contingently associated with a particular Way of Referring at a particular time without making much difference to Evans's point in the passage in question.

² For brevity, I omit numerous details of what such a theory of meaning would look like. An example, however, would be the kind of theory made popular by Davidson, based on a Tarskian truth theory.

- (1) The reference of 'Hesperus' = Hesperus
- (2) The reference of 'Hesperus' = Phosphorus

If it is a matter of indifference which of these two clauses is accepted then the theory in question is acceptable as a theory of reference but tells us nothing about the senses of the expressions in the language. A theory of reference of this kind can, however, be used simultaneously as a theory of sense (and thus a fully-fledged theory of meaning) provided that the clauses which assign reference to the terms of the language do so in a way which makes the sense of the term manifest. This can be done because to grasp a sense (at least on Evans's and Dummett's reading of Frege) is to think of the reference in a particular way and is therefore not something which could be stated other than in the process of stating the reference. The idea is that the senses of expressions are shown by the particular formulation of the theory of reference 'which identifies the references of expressions in the way in which one must identify them in order to understand the language' (Evans 1981, p. 282). One cannot simply state the sense of a term; instead one states what the reference of the term is, but the way in which one does this shows or displays its sense. This means that the formulation must include only clauses such as (1) and not clauses such as (2). In order to be made to understand (1) one must be made to think of the reference in the appropriate way; having done so, and having understood that the word 'Hesperus' refers to the object thus thought of, one has grasped the sense of 'Hesperus'.

The question, for our current purposes, is what is involved in identifying the references 'in the way in which one must identify them in order to understand the language'. This could mean that one must think of the reference using a particular Way of Referring in identifying it. But I can see no reason to insist on this, and indeed McDowell seems explicitly to reject such ideas:

Would knowledge that 'Hesperus' stands for Hesperus - in the context of suitable knowledge about other expressions, and suitable knowledge about the forces with which utterances may be issued - suffice for understanding utterances containing 'Hesperus'? Or would one require, rather, knowledge to the effect that the bearer of the name may be recognised or identified thus and so? Now patently this second, stronger requirement, interpreted in any ordinary way, insists on more than would suffice; for it insists on more than, in some cases, does suffice (1998, p. 178).

He then goes on to reject even the notion that the supposedly necessary knowledge could be possessed tacitly; the point is not merely that one does not *know*, explicitly, how one is to recognise or identify the reference, but rather that there is no particular way in which one has to do this in order to grasp the sense of the term. There is nothing about the way in which the reference is identified of which one needs to have even tacit knowledge in order to grasp the sense.

Whether or not Evans construed sense in the same 'austere' way is not absolutely clear, and I shall not attempt to find any further textual evidence on the matter; but it seems to me that he certainly could have construed sense in this way, without having to change much of what he said about reference or Russell's Principle. In any case, the apparatus of the austere theory of sense can be used to help make clearer the way in which I think concepts should be individuated. I do not think, in fact, that there is very much more to the individuation of concepts than the individuation of the words which are used to express them (or which could be used to express them, in the case of concepts which are never expressed). McDowell seems to have a similar idea:

Difference in sense between 'Hesperus' and 'Phosphorus' lies in the fact that the clauses in the theory of sense that specify the object presented by the names are constrained to present it in the ways in which the respective names present it. They meet this constraint - surely infallibly - by actually using the respective names. But it takes subtlety to find the metaphor

thus already applicable; it can easily seem to necessitate something more like a description theory of names (1998, p. 181).

What I take McDowell to mean - and what, in any case, I shall argue - is that once I have learned to use the name 'Hesperus' I have thus already acquired its sense, provided I know what it refers to. Thinking of the reference in the appropriate way simply means thinking of it qua Hesperus. Now, in this chapter I am still considering only the unshared thoughts of an individual so strictly speaking situations in which the sense of a name is learned from someone who already knows it are not our present concern. Shared thoughts, and the way in which the ideas currently being described can be used to provide an account of them, will be dealt with in the next chapter. But, of course, a theory of meaning of the kind described above can still be given for an individual's idiolect including concepts expressed by terms which are not shared with others.³

When clauses like (1) are part of a theory of meaning for an individual they should be understood as displaying the senses of the terms which the speaker could use in expressing thoughts; and thus displaying the particular concepts which that individual possesses. The individual's grasp of the concept consists in knowledge of a clause like (1); but it should be remembered that in order to grasp (1) it is necessary to be thinking of Hesperus, not merely of the word 'Hesperus'. One must be able to identify the object in question (but not by any particular means) and one must think of it qua Hesperus; one might be said to think of it in a 'Hesperus' way, and no further description of the 'way' in which it is thought of is necessary.

Further clarification of this general idea can be obtained by considering some views which David Kaplan developed after the publication of his *Demonstratives* (1989).

³ Cf. thinking about 'that cat' in private internal monologue (see above, section 2.5). I shall speak freely of the 'meaning of a term' even when this 'term' is left entirely unexpressed; what I am referring to is a term which would be used to express a given concept, if it were ever expressed. My view will not, however, rest on it being possible to make sense of the notion of an unexpressed term.

Although I cited Kaplan's content/character distinction combined with his 'dthat' operator above as an example of a version of Inferential Role Semantics (with which I disagree) Kaplan himself was never entirely convinced that he had found the correct account of cognitive significance and, therefore, of concepts. Recall his worry about proper names, for which in *Demonstratives* he assumed a causal chain theory of reference:

If the character and content of proper name words is as I have described it (according to the causal theory), then the informativeness of $[\alpha = \beta]$, with α and β proper names, is not accounted for in terms of differences in either content or character. The problem is that proper names do not seem to fit into the whole semantical and epistemological scheme as I have developed it. I claimed that a competent speaker knows the character of words. This suggests (even if it does not imply) that if two proper names have the same character, the competent speaker knows that. But he doesn't. (1989, pp. 562-3).

The problem is that for proper names the character, which assigns a content to the term given the context, simply assigns a reference to the name in the same way for all proper names referring to a given object. Consequently there are no differences in character to explain the informativeness of identities between coreferential names. This is a problem for Kaplan, given that character was supposed to correspond to the linguistic meaning of the term, and consequently to the mode of presentation expressed by the term. The same problem should occur for any context independent term on Kaplan's theory, not just proper names; so at best his theory works only for indexicals.⁴

Kaplan has subsequently had a rethink of these issues. In his *Afterthoughts* (1989a) he observes that in earlier work he had been trying to follow Frege in using 'a strictly semantical concept (character), needed for other semantical purposes, to try to capture

his idea of cognitive value' (p. 597). This assumes (probably correctly) that Frege had intended sense to serve the semantical purpose of determining reference; the idea being that something had to determine reference, and it would lead to a very neat theory if that same thing simultaneously determined cognitive significance. I have, of course, been arguing against this idea. Having discovered problems with this traditionally Fregean way of doing things, however, Kaplan hit upon a rather straightforward way of avoiding it:

The syntactic properties of 'Hesperus' and 'Phosphorus', for example, their distinctness as *words*, are surer components of cognition than any purported semantic values, whether objectual or descriptive.

If words are properly individuated, by their world histories rather than by their sound or spelling, a name might almost serve as its own Fregean *Sinn*. The linguistic difference between 'Hesperus' and 'Phosphorus' - the simple difference between thinking of Venus *qua Hesperus* and thinking of it *qua Phosphorus* - may be all the difference in mode of presentation one needs in order to be able to derive the benefits of sense and denotation theory. Words are undoubtedly denizens of cognition. If, through their history, they also provide the worldly link that determines the referent, then except for serving as content, they do all that Fregean *Sinn* is charged with. But they do it off-the-record, transparently and nondescriptively (1989a, p. 599).

Kaplan's idea is in some ways similar to Fodor's but offers a way of individuating concepts in situations which Fodor's account does not address. Although Fodor is an advocate of informational semantics he is still prepared to grant the neo-Fregean the notion of a mode of presentation. His official position is that he leaves it open what modes of presentation are so long as they are whatever distinguishes distinct but coreferential concepts such as WATER and H₂O (Fodor 1998, p. 15). In fact, however,

⁴ In fact I shall show below (chapter 6) that at the level of thought (rather than purely linguistic meaning) Kaplan's theory does not work for indexicals either.

it seems pretty clear that he thinks that what makes WATER and H₂O different concepts (assuming, for the moment, that H₂O is unstructured) is that they are associated with different mental symbols, that is, syntactically (and thus causally) different mental items with the same semantic properties. If concepts have their semantic properties essentially, however, they cannot be individuated purely by a physical entity inside the head, for such an entity has no intrinsic semantic properties.⁵ Moreover, individuating concepts by mental symbols alone would leave us with no way to answer questions about whether different people share the same concepts or whether, for example, a re-occurrence of the same symbol at a later time counts as a re-occurrence of the same concept. The latter problem is particularly pressing given the diachronic features of rationality for which we have been arguing.

Kaplan's account solves these difficulties by arguing, in effect, that for any concept there is a word which could be used to express the concept. He then ties the individuation of the concept to the individuation of the word. Thinking of something in the Way of Thinking associated with 'Hesperus' then just means thinking of it qua Hesperus.⁶ This is surprisingly close to McDowell's austere theory (surprising, that is, given the extent to which their views had previously differed).

Kaplan (1990a) provides a detailed account of how words are individuated. He contrasts the more traditional *orthographic* conception of words with his preferred account, the *common currency* conception. According to the orthographic conception occurrences of words in utterances, inscriptions and the like are tokens of a word type. Kaplan describes this as the 'typesetter's conception' of a word (1990a, p. 98). There can, however, be great differences in the way a word is spelled or pronounced in different dialects without our intuitions saying that it is not the same word. 'Colour' can be spelled 'color', schedule can be pronounced either 'shedje-yule' or 'skedje-oo-

⁵ Contra early Fodor (1980, 1987), but not necessarily contra the later, externalist Fodor (1994, 1998).

⁶ See also Ackerman 1979 for a similar view.

ul', and so on. There are well-known differences in regional and national accents. Yet we still consider the words produced to be occurrences of the same word.

Kaplan feels that these differences are not comfortably explained by regarding all these disparate utterances and inscriptions as tokens of the same type. A better account, he argues, can be obtained by the common currency conception according to which utterances and inscriptions are regarded as '*stages* of words, which are the *continuants* made up of these interpersonal stages along with some more mysterious *intrapersonal stages*' (1990a, p. 98). The account has certain (limited) similarities to the tentative 'causal chain' theory put forward by Kripke in *Naming and Necessity* (1972/1980). At some moment in time someone decides to dub a certain thing with a certain word. At this moment a new word is created. The continuity of the word through time depends only on the intention of the speaker to continue to produce the same word and similarly the continuity of the word from one person to another depends only on the intention of each speaker to utter the word that was uttered to them.

This means that when I learn a word by hearing someone uttering it, so long as I intend to utter the word I have learned, the sound (or inscription) that I make will count as an utterance of that word even if it sounds quite different from what I heard. My pronunciation may be different or my spelling may be different, but it is still the same word that I have reproduced. As Kaplan puts it:

No matter how poor the subject's *imitative* ability (his ability to make his output resemble his input), we can imagine circumstances in which we would say 'Yes, he is repeating that name; he is saying it in the best way that he can' (1990a, p. 104).

Kaplan thus construes words as worldly entities whose form, like that of other worldly entities, can change through time without their ceasing to exist. They continue to exist just so long as people continue in the practice of using them.

This leads to an individuation of words according to which, for example, different people called 'David' have different names; what they have in common is what Kaplan calls a 'generic' name. That is why it is possible to utter, for example, 'David is tall but David is not tall' without contradiction - one can be using two different words in such a sentence even though they are both examples of the same generic name. Many problems relating to cognitive significance are solved in this way, and the account seems to me to be perfectly compatible with conceptual atomism because the reference is simply thought of 'qua Hesperus', 'qua David' etc.

I shall not accept Kaplan's account without qualification, but I shall nevertheless adopt some quite similar ideas about how concepts can be individuated. The emphasis of my account is more on the individuation of concepts than of words. It may be, in fact, that there is a greater degree of separability between words and the concepts which they can be used to express than Kaplan allows. One reason for thinking this is that demonstrative and indexical terms cannot always express the same concepts (otherwise it would be impossible to claim that '*that* is not identical with *that*'). Yet it is not very intuitively appealing to suggest that every new utterance of 'that' expresses a new word. Moreover, I shall argue that the same concept can sometimes be expressed using different indexical terms at different times.⁷

One final point. Although Kaplan's observations about the differences in the spelling and pronunciation of a given word certainly show that words cannot be individuated by their precise spelling or pronunciation, he is not necessarily correct to conclude that *any* spelling and *any* pronunciation could count as an occurrence of the word. Word individuation could be comparable to doorknob individuation; it could be mind-dependent.⁸ What counts as a successful utterance of a word may be determined precisely by what we (typical members of the relevant linguistic community) are disposed to recognise as an utterance of that word. Someone who cannot pronounce

⁷ See chapters 5 and 6.

⁸ See above, section 2.1, for Fodor's discussion of mind-dependent concepts such as DOORKNOB.

the word in a way that we can understand simply cannot pronounce that word, whatever their intentions. This need not, however, preclude them from possessing the concept expressed by the word.⁹

3.2 *Concepts as referential episodes*

The correct way to individuate concepts, I suggest, is by one-to-one correspondence with *referential episodes*. The notion of a referential episode is a way of capturing the relevance of transparency to the individuation of concepts as implied by the Transparency Principle. When a person first refers to an object (or property) a referential episode begins. A new concept is thus created. The referential episode continues so long as there is diachronically transparent reference to the object; in other words, so long as the subject continues to possess a Way of Referring to the object and continues to think of it in a diachronically transparent way. If the subject then loses the ability to identify the object (by losing track of it perceptually or by forgetting a description which it uniquely fits, for example) the referential episode ceases and the concept ceases to exist. A referential episode begins, for example, when an object is first seen and identified demonstratively and it continues so long as the object is kept track of. If the object is lost track of, and the subject has no other means by which to identify it, the referential episode comes to an end. If the subject had acquired an alternative Way of Referring to the object through the perceptual encounter, however, the referential episode could continue when the object was no longer seen.

The notion of a referential episode thus cuts across particular Ways of Referring. A variety of different Ways of Referring may be used within a given referential episode,

⁹ Cf. Cappelen 1999. Kaplan's account is also criticised by Gregory McCulloch (1991) on the grounds that it cannot handle some of the problem cases discussed by Kripke in 'A Puzzle About Belief' (1979). I argue below (chapter 4), however, that such cases can be dealt with once it is correctly understood what is involved in sharing a concept with other people.

so long as the ability to identify the same reference is retained throughout, coupled with the necessary transparency. In order for there to be this transparency there will generally have to be a degree of temporal overlap when one Way of Referring is substituted for another. Imagine, for example, that at some point during a referential episode the reference is identified demonstratively. In order to continue the referential episode after perception of the object has ceased a recognitional or descriptive Way of Referring is required. But it is crucial that this should be acquired before the reference is lost track of perceptually; otherwise a question could arise in the subject's mind as to whether the same object was referred to at the different times. It is also crucial that the different Ways of Referring are understood by the subject to identify the same reference. A temporal overlap between the possession of different Ways of Referring to the same thing would not automatically constitute a single referential episode.

An analogy may be helpful here. A useful way of thinking of the difference between subject terms and predicates is to regard the subject term as a dossier of information about something and the predicate as an item of information in the dossier.¹⁰ This way of thinking of subject terms, in particular, provides a useful model for the dynamics of a simple concept. At some point in time a concept is created, for example due to a perceptual encounter with an object. Information about the object (that it has certain properties etc.) may be acquired over a period of time and added to the dossier of information which corresponds to the concept. Information may also be removed from the dossier or may simply become lost (due to a lapse of memory). The predicates in the dossier will often contain concepts which in turn have their own dossiers; the dossier for 'Jones' may, for example, contain the predicate 'is taller than Smith', and there may be a dossier for 'Smith' containing various other predicates, and so on.

We can broaden the notion of a dossier to include more than descriptive information. Picture a dossier for a given concept as being like a personnel file in an

office. For each employee there is, for example, a cardboard folder containing information relating to that employee. Corresponding to descriptive information in a dossier there are various papers in the file. But on the broadened notion of a dossier we can also incorporate a recognitional capacity; in our analogy we could think of this as, for example, a photograph of the employee placed in the file.¹¹ Demonstrative Ways of Referring are a little harder to fit into the analogy, but perhaps we could imagine a new employee coming into the personnel office for the first time. In order to start a new file we simply get the employee to hold the cardboard folder until some identifying information is acquired (a picture, some descriptive paperwork etc.) to put in it. If the employee leaves before this happens (if the reference is lost track of before another Way of Referring is acquired) then the file is simply discarded. But if the employee stays long enough we may be able to acquire identifying information about the employee to place in the file, in which case the file can be retained.

Now a system of files of this kind is normally operated by writing the name of an employee on each file and making sure that only items relating to the person with that name are placed in the file. But there is another possibility. Instead of writing names on the file we could simply allow the contents of the file to do the identifying. Imagine that within each file there is at least one item which uniquely identifies an employee - a photograph, a copy of a birth certificate, a paper with a precise job description written on it, and so on. These correspond to Ways of Referring. Each of these items performs a dual function: on one hand it identifies a person but on the other hand it also tells us something about that person. So a piece of paper with 'head of marketing' written on it tells us that the file belongs to the head of marketing, but it also tells us that the person in question *is* the head of marketing.

¹⁰ The notion of a dossier has been used by numerous philosophers. See e.g. Grice 1969, Evans 1982, Recanati 1993. The following account also has some similarities with the account of names given by Evans 1973.

¹¹ This analogy should not be taken too literally; I am not suggesting that a recognitional capacity works by having a mental image of the reference which is compared with what is perceived.

If there was only one item in the file this would, of course, be uninformative. But now suppose that the head of marketing comes into the personnel office; this corresponds, in our analogy, to a perceptual encounter with a referent. Further information can be acquired from the head of marketing. First of all, we must establish who the person in the office is. If we discover (for example through the testimony of the person in question) that this person the head of marketing we can look through the files and find that there is just one file with this information in it (this corresponds to the identification of a perceptually presented object as one which was previously only identified by description). So any further information that we can acquire from the person now present can be put in this file. We discover the person's name (Smith), room number, National Insurance number, and so on. A photograph could also be added. Some of the information might not be uniquely identifying (e.g. 'used to work at Head Office'), but some of it might be.

So now when Smith leaves the room we have a file containing many ways of identifying Smith. But it is still the same file. The original piece of paper which said 'head of marketing' could now be lost - it might, for example, fall out of the file by accident (corresponding to a lapse of memory). But the same file would be retained, because it would still retain the means to identify its subject. In fact, over the course of time, there may be many replacements of items in the file. It could be the case that nothing ever stayed in it permanently. Yet even though there was nothing about the file itself (the cardboard folder) to determine who it was about it would still retain its identity as a particular file about a particular person. This corresponds to the way in which a referential episode continues through a temporal overlap of Ways of Referring.

If the file was ever emptied completely it would lose its identity, of course. It would simply cease to exist. There was really no need to include the cardboard folder in the story at all - the same system could be operated with piles of paper on a table. When no papers are left, the pile vanishes. But a new pile can be created with any items which uniquely identify an individual. The metaphysics of a pile of papers in which

items can be added or taken away thus mirrors certain aspects of the metaphysics of a referential episode, and thus of a concept.¹²

A second file could, of course, be opened in error for the same individual. Suppose that in the personnel office we receive a phone call from the deputy director. Unaware that there was such a person, we open a new file using a piece of paper with 'deputy director' written on it. As time passes, other items are added to the deputy director's file. Yet all along the deputy director was none other than Smith, who holds both the post of deputy director and head of marketing. This, of course, corresponds to a case in which there are coreferential terms which are not known to be coreferential.

Finally we must consider what happens if false information is placed in the file. This could lead to a file containing information which identifies two different individuals. Either we say that the file is about Smith but contains false information actually relating to Jones, or vice versa, or we say that the file fails to be about any particular individual. Which of these is correct depends on which of the identifications is taken to be non-negotiable. Typically at any moment in time there will be some means of identification which is regarded as non-negotiable while there may be others for which doubts can be entertained. I may recall that the file was started as belonging to the head of marketing and may therefore treat a paper which says 'chief accountant' as rogue information when I discover that the head of marketing is not the chief accountant. Or I may have forgotten how the file started out and I may come to treat 'chief accountant' as non-negotiable and the paper which says 'head of marketing' may therefore have become the rogue information instead.

A Way of Referring which is non-negotiable at one time might lose this status at some later time without changing the concept. In order for a file to retain its identity, however, it must continue to be about the same person. That means that all new information which is treated as non-negotiable must *correctly* identify the same individual. In terms of concepts, this corresponds to the fact that the same concept is

¹² Consequently, contra Frege, concepts are not timeless entities on this view.

only retained if its reference is preserved. In those rare cases where there are two simultaneous Ways of Referring which pick out different referents but are both treated by the subject as non-negotiable then there is simply no unambiguous concept possessed. The subject is in a confused state and no unambiguous thought is possible.¹³

Despite what has been said so far, there are situations in which it might appear that a concept can be retained through changes of Ways of Referring even though there is no strict overlap between the different Ways of Referring. It depends, however, on how Ways of Referring are individuated. Consider, for example, an ongoing demonstrative identification of an object over a period of time. Clearly there are circumstances in which different occurrences of the word 'that' at the same time express different concepts even if both tokens of 'that' refer to the same thing. Thus one might, for example, see each end of a long ship through two different windows and think 'that is a steamer but that is not a steamer', without contradiction.¹⁴ Consequently it will not suffice to claim that there is only a generic 'demonstrative Way of Referring' which is the same for all demonstratives.

Now, if we consider only how things stand at a given moment in time it might seem sufficient to individuate demonstrative Ways of Referring in terms of directions: the information from one end of the ship is received from a certain direction while the

¹³ Note the related comments about unnoticed switching of the reference during an episode of demonstrative thought in chapter 1, above. The idea that certain Ways of Referring are held as non-negotiable at certain times may hold a clue to the solution of the following rather vexing problem in metaphysics. Suppose that an object can be identified in virtue of two distinguishing features but the object then splits in two, with one of the previously distinguishing features belonging to each half. Since the two products of the fission are not identical to one another they cannot both be identical to the original object so it may seem that neither of them is identical to it; the original object has ceased to exist. Yet sometimes we say that an object continues to exist when a piece breaks off it. So how do we determine which, if either, of the fission products is the original object? Perhaps this depends on which identifying feature is held as non-negotiable; if just one is held as non-negotiable, then the object survives and is retained as the half with the distinguishing feature (and now lacks the other property). But if both features were held as non-negotiable then the original object has ceased to exist. This way of approaching the problem does rather assume that identify is in the eye of the beholder. But as anyone who has studied the metaphysics of identity over time knows, cases like this present an intractable problem for almost all theories of identity which approach the problem in mind-independent terms. See surveys of the literature on this subject see e.g. Oderberg 1993, Gallois 1998 chapter 1, Sider 2000.

information from the other end of the ship is received from a different direction. But over a period of time the same object may move around relative to the observer while the observer keeps track of it. We could, of course, describe this as an infinite series of instantaneous non-overlapping Ways of Referring, where Ways of Referring are individuated in terms of directions. These non-overlapping Ways of Referring would be associated with a single Way of Thinking. But this seems an unnecessarily clumsy way of putting things. It seems more reasonable to say that human beings have an ability to keep track of moving objects and it is the ongoing exercise of this single ability over a particular episode that is relevant to individuating the Way of Referring.¹⁵ Different simultaneous demonstrative Ways of Referring could then be distinguished from one another by, for example, the different information channels upon which they depend at a given time.

A similar situation occurs with certain other types of Way of Referring. Consider, for example, keeping track of a place while moving around with one's eyes closed and without using any other sensory modality.¹⁶ One could still keep track in a very limited way provided one knew roughly how one was moving. One might, for example, have a way of moving slightly to the left. Provided one had the necessary understanding of spatial relations one would then know that the place which was formerly 'here' would now be 'slightly to the right'. Given that one had no sensory input the place could only be identified egocentrically as, for instance, 'here' or 'slightly to the right'.¹⁷ These

¹⁴ This example has been used by numerous authors but, as far as I know, is originally from Perry 1977, p. 483.

¹⁵ Ruth Millikan (1993, p. 339) conspicuously misses this point in arguing that modes of presentation cannot be semantic (i.e. involving an intrinsic connection with the reference) rather than merely 'psychological'. She assumes that wearing left-right inverting prisms in a situation like the one described would reverse the matching up of modes of presentation to referents. Dynamic considerations show that this is simply the wrong way to individuate modes of presentation. Further arguments for the 'dynamic' view of demonstratives will be given in chapter 5, but it should be sufficient to note that the same arguments about the diachronic nature of rationality given in chapter 2 apply a fortiori here. This is another straightforward application of the Transparency Principle.

¹⁶ One normally uses one's senses to keep track of places just as one does when keeping track of objects; the way of keeping track described here is a special case and could only work in a limited way (see below, chapters 5 and 6). Something very similar can occur in keeping track of time.

¹⁷ See below, section 6.6, for more on egocentric and other ways of identifying places.

identifications could not overlap, yet one could still think of the place in the same way throughout.

One final example. Imagine a room containing several moving objects which are exactly alike except for their colours. These change gradually and continuously in a regular way such that there are never any two of the same colour at the same time. Someone who observed this and then left the room could continue, at least for a limited period, to think about one of the objects by using its colour to identify it. But this would require a continuous updating of the Way of Referring in order to keep up with the changes in the colour of the object. Once again we have a choice between describing this as a single Way of Referring or an infinite series of different non-overlapping Ways of Referring. I think the former description is the more natural, but if we chose the latter description it would no longer be correct to insist that Ways of Referring must always overlap within a referential episode. The difference, however, is purely terminological.

3.3 Analyticity and holism

One advantage the version of Conceptual Atomism for which I have been arguing is that it makes it easy to explain away the admittedly strong intuitions most of us have that sentences like 'bachelors are unmarried men' are analytic. It is desirable to be able to explain this because Conceptual Atomism is inconsistent with analyticity. It would be inconsistent with Conceptual Atomism, for example, to claim that it is constitutive of the concept BACHELOR that possession of it requires possession of the concepts UNMARRIED and MAN.

In a nutshell, intuitions of analyticity can be explained away because of the fact that there are some concepts for which there are limitations on the ways in which it possible

to identify the reference. Thus, for example, if it is only possible to identify bachelors in virtue of their being unmarried men then anyone who possesses the concept BACHELOR will inevitably find the statement 'bachelors are unmarried men' compelling. But it does not follow from this that an account of the individuation of the concept BACHELOR makes essential reference to the concepts UNMARRIED or MAN; it can be individuated in terms of its reference and the circumstances of its creation and continuity (that is, in terms of a referential episode), without any need to mention the specifics of the way in which the subject identifies the reference.¹⁸

On this issue I entirely agree with Fodor (1998, pp. 80-86). Fodor expresses his view as a modification of Putnam's (1983) view that there are 'one-criterion' concepts. Despite being a fan of Quine's (1961) attack on the analytic/synthetic distinction, Putnam was bothered by the fact that on Quine's view the only thing that makes some sentences harder to deny than others is that they are more 'central'. But certain beliefs, such as 'bachelors are unmarried men' seem just impossible to deny in a way that a relatively central belief like Newton's second law, ' $F = MA$ ', is not. If beliefs of the former kind are analytic, however, this will be of far less philosophical interest than, for instance, ' $F = MA$ ' being analytic. So, according to Putnam, much of what is philosophically important in Quine's denial of the analytic/synthetic distinction still applies even if there are a few philosophically uninteresting analytic sentences like 'bachelors are unmarried men'.

Putnam's claim, then, is that although there are sentences which are analytic they are among the least central ones; their truth or falsity has virtually no consequences for the truth or falsity of other beliefs (unlike central ones such as ' $F = MA$ '). What makes these sentences analytic is that they contain a one-criterion concept: a concept for

¹⁸ Although analyticity conflicts with Conceptual Atomism it should not be assumed that advocates of Inferential Role Semantics are invariably committed to analyticity. Peacocke (1997, p. 234), for example, suggests that the notion of the canonical commitments of a concept can explain why certain sentences are a priori without those sentences being analytic. This is a similar idea to the one being advocated here. My difference with Peacocke concerns only his claim that canonical commitments are constitutive of a concept.

which there is only one criterion which can be used to tell whether or not the concept applies. So, for example, there may be only one way to find out whether something is a bachelor and that is to determine whether it is an unmarried man. Hence BACHELOR is a one-criterion concept and that is why, according to Putnam, 'bachelors are unmarried men' is analytic.

Putnam could presumably have said something similar about 'recognitional' concepts (as discussed and rejected in chapter 2). Although these do not really give rise to analytic sentences they are supposed to have just one criterion of application (an experience of a certain kind). So given the experience it would be something akin to analytic that the concept applied.¹⁹ Thus it could be claimed that it is constitutive of a recognitional concept that it is related to certain kinds of experiences via a recognitional capacity just as it would be constitutive of BACHELOR that it stands in certain relations to the concepts UNMARRIED and MAN.

As Fodor points out, however, Putnam's account is circular. The problem is essentially that Putnam attempts to explain what analyticity is in terms of one-criterionhood, but it is impossible to explain one-criterionhood without appealing to analyticity. For in order to determine that something is a one-criterion concept it is necessary to have a way to count criteria. It is necessary to be able to determine, for example, whether 'unmarried man' and 'man who is not married' count as the same criterion. But whether they do depends on whether they are synonymous, and synonymy is interdefinable with analyticity. So asserting that BACHELOR is a one-criterion concept is really not much different from just asserting that 'bachelors are unmarried men' is analytic, and it cannot therefore explain the supposed analyticity of the latter sentence.²⁰

¹⁹ Perhaps not strictly analytic, though, because the fact that I am experiencing a certain object as looking red, for example, is not sufficient for its being red (whereas by contrast Jones's being an unmarried man is sufficient for Jones being a bachelor). As suggested in chapter 2, there is always an element of theory behind even the most basic recognitional judgments; they always involve a hypothesis rather than an inference.

²⁰ Fodor (1998, p. 82) attributes this argument to Jerry Katz.

The intuitive idea behind one-criterion concepts can, however, be used to explain the spurious intuition of analyticity (as opposed to accounting for a genuine property constitutive of certain concepts, as Putnam tries to do). To put Fodor's point in my terminology, whereas for most concepts there are many possible Ways of Referring there are some concepts for which there are very few. Whether or not we say that for some concepts there is precisely one Way of Referring does not really matter. What matters is that for such concepts certain truths seem undeniable simply because it is only through those truths that we have any way to pick out the reference of the concept. The only way to tell whether someone is a bachelor, for example, is to determine whether that person is an unmarried man. Hence the strength of the intuition of analyticity. But that does not show that BACHELOR is *constituted* in terms of UNMARRIED and MAN or that no account can be given of the former without mentioning the latter. To think otherwise would be to conflate the epistemic property of being a priori with the semantic property of being analytic. The same could probably be said about the supposed recognitional concepts: there may as a matter of fact be only one way to determine whether something is red but that does not make the recognitional capacity constitutive of the concept.

In any case, there is a further reason to reject the notion of one-criterionhood as a means of capturing a notion of analyticity. The reason is that strictly speaking there can be no concepts the identification of whose reference requires the possession of specific concepts. For if, as I am arguing, concepts are individuated by referential episodes then there are no particular concepts which one must possess in order to possess the concept BACHELOR, even though bachelors can only be identified by virtue of being unmarried men. For we must keep in mind that there can be different concepts which share the same reference. On the account of concepts for which I am arguing there can be a potentially infinite number of concepts referring to any given thing. Bachelors can be identified descriptively as unmarried men, or unmarried men', or unmarried men"...etc., where MAN, MAN', MAN",...etc. are all different concepts referring to men

(and the same applies, *mutatis mutandis*, for 'unmarried'). So although it may well be the case that in order to possess the concept BACHELOR it is necessary to be able to identify men it does not follow that in doing so it is necessary to think of men under any particular mode of presentation. The most that can be concluded is that in order to possess the concept BACHELOR it is necessary to possess a Way of Referring to men.

It might be objected that in some cases it is not possible to possess more than one concept with a given reference. Although there may be different ways of identifying men, and consequently one could simultaneously possess MAN and MAN' without realising that they referred to the same thing, there may be some concepts (PLACE might be an example) for which it would be impossible to possess two co-referring concepts without realising that they referred to the same thing. This objection would, of course, apply only to a very limited set of cases. But in any case it is question-begging for it ignores the possibility that different concepts (such as different place concepts) could be different because they corresponded to referential episodes that took place at different, non-overlapping times.

Analyticity is related to meaning holism. Holism is a large and complex subject; too complex, in fact, for it to be possible to give a comprehensive discussion of it here. I shall, however, make a few general remarks. Holism comes in many different versions. Some versions have the consequence that when one theory is replaced by another the terms of the theory change their meanings. So, for example, when one theory about electrons is replaced by another the word 'electron' changes its meaning - it no longer expresses the same concept. A typical argument for this would appeal to the idea that electrons are theoretical posits, unobservable entities posited to explain observable phenomena. What a different theory posits must be different; if the 'electrons' of the new theory have different properties then they are different entities.

Radical versions of this position (such as Kuhn's (1962) theory) claim that when there is a major change in a theory all of our concepts change. But weaker versions are also possible, in which it is only concepts which stand for unobservable posits like

electrons which change, thus avoiding the rather implausible consequence that concepts like COW and TREE are replaced by different concepts every time there is a discovery in sub-atomic physics. Cows and trees, after all, seem to be most obviously identifiable as things with which we are familiar from our encounters with them, and they are still the same entities when we make a new discovery about their nature. But, it might be argued, the same cannot be said about electrons; the only reason for thinking that they exist at all is that they play a certain specific role in a theory and if we reject the theory we reject the entities it posits.

Conceptual Atomism is radically opposed to holism but the version for which I have been arguing helps make it clear what is right and wrong in the arguments for holism. Whether or not the concept ELECTRON can be retained through a change of theory depends on whether or not the word 'electron' is expressive of a single referential episode. So long as electrons continue to be identified as, for instance, the particles transmitted in an electrical spark then there is no reason why the word 'electron' should not express the same concept on both the old and the new theory; the two theories just say different things about the same entities. If, on the other hand, there is not a continued referential episode then the two theories use different concepts for which they happen to use the same word. I assume that the latter scenario very rarely occurs.

Some philosophers who reject global holism nonetheless maintain that there are 'local' holisms. These are families of interdependent concepts such that possession of one member of the family requires the possession of others. PLACE and SPATIAL RELATION would be an example; it might be held that one cannot possess one of these concepts without possessing the other.²¹ Local holisms, however, can be treated by a Conceptual Atomist along the same lines as analyticity. Certain groups of concepts might be such that identifying the reference of one requires being able to identify the reference of another, and vice versa. But it does not follow from this that an account of

the individuation of one of these concepts makes essential reference to any of the others.

Some philosophers are very interested in investigating what one must be committed to in order to possess a given concept; the fact, for example, that no one could be said to possess the concept PLACE who did not have a certain grasp of the spatial relations that one place can stand in to another.²² Such matters are certainly important. But these studies should be understood as placing constraints only at the level of reference, not at the level of concept individuation. They concern the details of what is required in order to satisfy Russell's Principle for a given concept; but this is not part of an account of the individuation of the concept.

3.4 Psychology and causation

The theory which has been developed has the consequence that two people could be in identical internal physical states, be thinking about the same thing, and yet be entertaining numerically different thoughts because their concepts were associated with different referential episodes. This puts the theory in conflict with a familiar argument that says that psychological states, in order to be explanatory, must be individuated in terms of the causal powers of the states. The locus classicus for arguments of this kind is probably the earlier work of Fodor (1980, 1987) but similar views have been expressed by many philosophers. Consider, for example, this passage by David Lewis:

The main purpose of assigning objects of attitudes is, I take it, to characterise states of the head; to specify their causal roles with respect to behaviour, stimuli, and one another. If the assignment of objects depends partly on something besides the state of the head, it will not

²¹ Cf. Peacocke 1997, p. 243.

²² See for example Peacocke's notion of 'canonical commitments' as described above, section 2.3.

serve this purpose. The states it characterises will not be the occupants of the causal roles (Lewis 1979, p. 526).

On the face of it a theory which individuates concepts in terms of referential episodes would be vulnerable to such an objection if it were valid. The objection assumes, however, that a science should not admit different entities which do not differ in their causal powers. But this is obviously false; the correct formulation is that a science should not admit different *types* of entities which do not differ in their causal powers. There is only one type of particle called the electron but there are countless electrons. Physicists make progress by thinking about types of entities; they rarely, if ever, think about particular electrons. The same should be said of psychologists. Referential episodes, and hence concepts, are tokens of types of thought components. They are created by a particular person at a particular time. Psychologists make progress by thinking about types of concepts and need rarely concern themselves with particular concepts. But this does not render the notion of a particular concept, distinguished from others of the same type by the circumstances of its creation and continuity, in any way problematic.

APPENDIX A

Reference, Causation and Information

What does the version of Conceptual Atomism for which I have been arguing say about the causal and/or informational relation between a concept and its reference? In answering this question we must take account of two relatively popular theories. Firstly, the 'causal theory of reference' says that the semantic properties of a concept are determined by a causal chain leading back to an 'initial baptism' of the reference. This theory is well known from Kripke's tentative advocacy of it in *Naming and Necessity* (1972/1980). The second theory, often known as 'informational semantics', says that the semantic properties of concepts are determined by informational relations between the concept and its reference (where 'information' is understood in the technical sense found in the work of Claude Shannon (1948) and brought to the attention of philosophers by Fred Dretske's *Knowledge and the Flow of Information* (1981)). I shall argue, perhaps rather controversially, that no causal encounter between the subject and the reference is necessary for the creation of a concept and that although a concept stands in an informational relation to its reference this can occur without any information or causal influence having been received from the reference.

There is no doubt that in the enormous majority of cases a referential episode begins with a causal encounter with the reference. This takes the form of a perceptual encounter which provides the subject with a demonstrative Way of Referring. A demonstrative Way of Referring constitutively involves a causal relation to the reference; it would be impossible to identify something demonstratively without being causally influenced by it. This is not true of recognitional or descriptive Ways of Referring, of course; indeed, part of the point of such Ways of Referring is to make it possible to refer to something which one is not causally influenced by at the time. So

the question of whether a causal encounter with the reference is necessary for the creation of a concept is determined by whether or not it is possible to begin a referential episode with a recognitional or descriptive Way of Referring without having acquired either of these as a result of a causal/perceptual encounter with the reference.

There may, of course, be certain concepts for which the idea of causal influence is not appropriate; it is hard to see, for example, how one could acquire concepts like ISOTOPE or CONTRACT other than by using a definite description to fix the reference. So let us limit our attention to those cases in which causal influence seems most necessary: reference to observable physical objects and properties. If the creation of primitive concepts of this kind does not require causal influence from the reference then I shall assume that the same conclusion follows for concepts of all kinds.

The advocate of a causal theory of reference need not insist that the chain of causation goes all the way back to the reference (though the theory is often stated in this way). Indeed, Kripke's (1980, p. 96) original idea was just that there must be a chain of causation back to the 'initial baptism' of the term, and this initial baptism could be achieved by using a description to fix the reference just as easily as an ostension. In such a case there would be no chain of causation leading back to the reference.

It is necessary, however, to be able to distinguish the thoughts of someone who entertains a simple concept created using a description from those of someone who entertains a logically descriptive thought containing the same description (a thought involving a complex conceptual structure with existential quantification). Evans, for example, considered what he called 'descriptive names'; names which are brought into existence by stipulating that they stand for the reference of a certain description. These, according to Evans, differ from ordinary proper names, which are associated with no particular description. Evans's example of a descriptive name is 'Julius', which is introduced as 'whoever invented the zip' (1982, §§1.7, 1.8, 2.3, 2.4, 2.5). He argued that descriptive names are referring terms but not 'Russellian' ones; that is, if nothing is

picked out by the description (i.e. if there is no reference) a sentence containing the name nonetheless expresses a thought. For example (1982, pp. 50-1), the thought expressed by:

(1) Julius is *F*

is the same one expressed by:

(2) The inventor of the zip is *F*

Evans holds this view even though the name and the description embed differently in modal contexts. Consider for example (1982, p. 60):

(3) If you had invented the zip, you would have been Julius

This sentence, as Evans observes, could only be correct on the assumption that 'Julius' is just an abbreviation for 'the inventor of the zip'. But Evans distinguishes the use of 'Julius' as an abbreviation from its use as a descriptive name. If 'Julius' is a descriptive name then sentence (3) cannot be an acceptable statement.

It is unclear, however, why Evans thought that it was legitimate to say that 'Julius' does not simply abbreviate 'the inventor of the zip' while implying that both the former and latter terms express the same mode of presentation (which presumably they must if the thought expressed by 'Julius is *F*' is the same one expressed by 'the inventor of the zip is *F*'). On the face of it the modal differences between the two terms imply a difference in the thoughts expressed using them. Evans's argument for identifying the modes of presentation (i.e. for identifying the thoughts expressed using (1) and (2)) is simply that:

Someone who understands and accepts the one sentence as true gets himself into exactly the same belief state as someone who accepts the other. Belief states are distinguished by the evidence which gives rise to them, and the expectations, behaviour, and further beliefs which may be derived from them (in conjunction with other beliefs); and in all these respects, the belief states associated with the two sentences are indistinguishable. We do not produce new thoughts (new beliefs) simply by a 'stroke of the pen' (in Grice's [1969, p. 140] phrase) - simply by introducing a name into the language (1982, p. 50).

But surely only the use of 'Julius' as an *abbreviation* for 'the inventor of the zip' is a stroke of the pen; introducing it as a name, by contrast, marks a logical difference manifested by the differences in modal embedding. Names simply contribute their references to the truth conditions of sentences containing them, whereas descriptions have a quite different logical structure. Evans did not deny this, yet he thought that sentences containing descriptive names expressed the very same thoughts as those containing the corresponding descriptions.

I suspect that Evans's view rests on the same error that is made when it is assumed that because bachelors can only be identified in virtue of being unmarried men it follows that 'bachelors are unmarried men' is analytic. Evans might have said: 'someone who believes that "there were three bachelors at the party" gets into exactly the same belief state as someone who believes that "there were three unmarried men at the party"'. But we have seen that this is not so; although the belief states have many similarities they do not involve the same concepts. Someone entertaining a descriptive name associates only one Way of Referring with it (though more can subsequently be acquired). Consequently such a person is in a position to make an a priori assertion that 'Julius was the inventor of the zip', but this does not make the sentence analytic. So the intuition that (1) and (2) express the same thought can be explained away in the same way as intuitions of analyticity.

I argued above that just because bachelors are identified by virtue of being unmarried men it does not follow that possessing the concept BACHELOR requires one to possess any particular concepts (at most, one needs *a* concept whose reference is the property of being male and another whose reference is the property of being unmarried). With regard to descriptive names it is worth emphasising this notion of identifying the reference *by virtue* of something. Thinking about something using a perceptual demonstrative Way of Referring, for example, involves identifying something at least partly by virtue of the fact that the item in question is the causal source of the current perception. It could be argued, moreover, that anyone who has even quite an elementary understanding of perception must believe that this is what is going on when something is identified demonstratively. Certainly one could not coherently see an object and think ‘that is *not* causing my current perception’.

It would be quite a different matter, however, to claim that demonstrative thought involves identifying the object *as* the causal source of the current perception; that is, thinking of it as whatever is referred to by the description ‘the causal source of my current perception’. This would entirely misrepresent the nature of the thought; the thought that ‘that is *F*’ is of simple subject-predicate form and does not involve existential quantification. It is certainly not the same as the thought that ‘the causal source of my current perception is *F*’.

Now, the thought expressed by (2) does involve existential quantification (assuming a Russellian analysis of descriptions, at least). If it expressed the same thought as (1) then (1) would also involve existential quantification. But it is not obvious that thinking of Julius *by virtue* of the fact that Julius was the inventor of the zip involves existential quantification at all. For there are two possible ways to analyse a sentence like ‘Julius is the inventor of the zip’.¹ Using ‘ ϕ ’ for ‘invented the zip’ the two analyses are:

¹ Cf. examples (15) and (16) in section 2.5, above.

(4) $\exists x [x = \text{Julius} \ \& \ \phi x \ \& \ \forall y (\phi y \rightarrow y = x)]$

(5) $\phi \text{Julius} \ \& \ \forall y (\phi y \rightarrow y = \text{Julius})$

I can see no reason to insist that (4) is the correct analysis. On the contrary, when one thinks, about a , that ' a is the ϕ ' one does not make an explicit assumption about existence any more than one does when one makes any judgment of the form ' a is F '. One merely thinks that a has the property of being ϕ , and nothing else does. Setting aside the possibility of a free logic one can of course derive an existentially quantified sentence from any simple subject-predicate sentence and indeed one can derive (4) from (5) in this way. But they do not express the same thought.

It might perhaps be objected that one cannot entertain a thought like (5) unless one has some independent means of identifying Julius. But we would not say this in the demonstrative case. It is not incoherent to think '*that* is the causal source of my current perception' (or perhaps '*this*' perception); it does not require '*that*' to be identified in some independent way. In both cases what is thought is uninformative, but not incoherent. And if another Way of Referring to the same object was acquired the thought need no longer be uninformative.

So my suggestion is that provided we keep in mind the distinction between thinking of something *by virtue* of its having a certain property and thinking of it *as* 'whatever has that property' then it will no longer seem compelling to assume that (1) and (2) express the same thought. It follows that a new referential episode can begin, and a new concept can be created, even when the thinker has never been causally influenced by the reference because the initial Way of Referring need not be demonstrative.² If, as a matter of fact, nothing is picked out by an initial descriptive Way of Referring (such as 'the inventor of the zip') then the situation can be thought of as analogous to a case of hallucination in which nothing is referred to by a demonstrative. It may be that the

fact that an existentially quantified thought like (4) is always derivable from a thought like (5) could be of some help in making sense of what a person in such a situation *does* think, but I shall not pursue this here.

François Recanati (1993, pp. 109-112), who agrees with much of what Evans says about descriptive names, makes an interesting claim about the way in which a descriptive name can become an ordinary name. Recanati claims that a descriptive name is 'created in anticipation of a time at which it will be possible to think of its reference non-descriptively' (1993, p. 109). His idea is that a descriptive name is introduced via a reference fixing stipulation, then at some later time the thinker 'makes contact' with the reference, recognising it because it satisfies the reference fixing description. It is then possible to acquire further information about the reference and further means of identifying it, and thereafter it is even possible to forget or reject the original reference fixing description. When 'contact' is made with the reference the thinker receives information from it, and is only when information has been thus received that non-descriptive (*de re*) thought about the reference is possible. Recanati does, however, allow that someone who is only under the illusion of having received information can also think about the reference non-descriptively.

I do not agree with Recanati's view for the same reasons for which I disagree with Evans. The implied assumption that only a Way of Referring which involves causal interaction with the reference allows the creation of a simple, directly referring concept is arbitrary and unprincipled. But Recanati's discussion does bring out one interesting fact: that in order to think anything about Julius (other than that he invented the zip) it is normally necessary to at least believe oneself to have encountered Julius, either directly or indirectly (e.g. through the testimony of others). Be that as it may, this seems to be a psychological rather than a logical point. It may well be that in practice no one would ever take themselves to have *de re* beliefs about Julius without having

² This is consistent with Kaplan's (1978, p. 241) claim that *dthat* (the first child to be born in the twenty-first century) can refer directly (Kaplan was, of course, writing in the twentieth century).

been causally influenced by him, but that does not show that causation is a necessary condition for the availability of the concept expressed by 'Julius'.

An example that helps bring this out is Davidson's (1987) 'swampman'. The swampman is an atom-for-atom replica of Davidson, created by sheer chance out of stray molecules in the air moments after Davidson himself is vaporised by a bolt of lightning in the same place. Despite the fact that the swampman's brain is in the same state and in the same environment that Davidson's had been in, Davidson claims that the swampman has no thoughts at first because his words (and presumably concepts) were not learned; they have no causal history. Now, I have rejected this causal claim.³ Many of the swampman's beliefs may be false, especially those involving memory. But the swampman will still have Ways of Referring to many objects and properties which do not require any causal connections with the reference. Moreover, the swampman will not take himself to be in the position of someone whose concepts have all been introduced by stipulation even though he has never actually been 'in contact' with anything. He will be in the position which Recanati thinks is necessary for non-descriptive thoughts but without ever having been causally influenced by anything.

Now let us turn our attention to the informational relation between a concept and its reference. Ever since Dretske published his book on the role of information in thought and knowledge (1981) a number of philosophers (including Fodor) have endorsed the idea, known as 'informational semantics', that what makes the concept *X* about *x* is that *X* carries information about *x*.⁴ A demonstrative utterance of 'there is an *x*', for example, carries the information that there is an *x* in the vicinity. In order for this to be so it is normally necessary for there to be a causal relation between appearances of *x*'s and tokenings of the concept *X*. So COW refers to cows, on this story, because tokenings of 'there is a cow' carry the information that there is a cow (and not

³ Davidson may have changed his mind about this too. He said at a conference in 1999 (in Aix-en-Provence) that he no longer believes what he said about the swampman, though he did not give details.

something else) in the vicinity. This can only happen, however, if it is cows (and not something else) that cause tokenings of 'there is a cow'.

The difficulties involved in establishing precisely which causal relations determine the informational relation which in turn determines the reference are considerable. But for those who accept Russell's Principle it is not necessary to solve these problems. Russell's Principle surely has informational semantics as a corollary; since, in order to possess the concept COW, I must know how to distinguish cows from all other things my 'cow'-tokenings must carry information about the presence of cows. It is this 'knowing which' requirement that sets up the informational relation. So it is a consequence of Russell's Principle that a concept must stand in an informational relation to its reference.

Consequently, in order to be consistent with what I have said about causation it is necessary to hold that the requisite informational relation can be established, at least in principle, prior to receiving any causal influence from the reference. Now, it is vital to distinguish the notion of *receiving* information from x from the notion of *'carrying'* information about x . By the former, I mean the transmission of information from one place to another. This usually requires causal influence to pass from one place to another.⁵ By the latter I mean that there is information in a certain item about some other item. This can occur without causal influence ever having been passed from one item to the other. The smoke alarm on my ceiling, for example, carries the information that the building is on fire when it sounds, but it also carries the information that the building is not on fire when it is silent. Since the building has never been on fire there has never been any relevant causal influence from a fire to the smoke alarm. But so

⁴ I shall assume that the notion of 'information' is familiar and I shall not give a detailed description of it here. Dretske's (1981) book contains an excellent exposition.

⁵ In the terminology used by Hans Reichenbach (1956) and Wesley Salmon (1984) to describe causation this can be thought of as the transmission of a 'mark' by a 'causal process'.

long as it is working properly it is nevertheless *fire*, specifically, that the smoke alarm's silence tells me about the absence of.⁶

So, by the same token, even if I have never encountered a cow it is possible in principle for me to possess the concept COW. I simply have to get into a state such that if there were any cows around I would be inclined to recognise them as such. My lack of 'cow'-tokenings then carries information about a lack of cows in the vicinity. It may, of course, be difficult to get into the requisite state (i.e. to satisfy Russell's Principle) for certain concepts without causal influence from the reference. If the reference can be identified descriptively then it is easy.⁷ In practice it may be harder for concepts which can only be identified non-descriptively. But it is not impossible.

Dretske, however, despite allowing the possibility of innate concepts, disagrees that for other concepts it is possible in practice to be in the requisite informational relation to the reference without being causally influenced by it. He says this because of counterexamples such as the following (from Dretske 1981, chapter 9). Suppose someone tries to teach me the concept RED by showing me a white object with red light shining on it. What I am shown is not really red, but I nevertheless acquire a recognitional capacity which will pick out things which are, as a matter of fact, red. Dretske argues that I do not thereby acquire a concept referring to the colour red because I have learned the concept in such a way that I am actually sensitised to things which seem red rather than things which are red. Only by being shown something which is really red, he argues, can I acquire the sensitivity required to possess the concept RED.

There is surely a flaw in this reasoning. Certainly if I am aware that I am looking at a white object with red light shining on it then I might get the wrong end of the stick

⁶ The lack of sound from the alarm does not imply a lack of information. In nuclear power stations in order to protect against faulty alarms they use alarms which work in reverse: they make a regular sound when there is no danger and it is only when they fall silent that there is danger. For these alarms it is the sound, rather than the silence, that carries the information about the absence of fire (or radiation etc.)

and think that 'red' refers to a heterogeneous collection of things which either are or look red. But I need not be aware of the circumstances. I could be tricked into thinking that I am looking at an object under normal lighting conditions, and I may make the assumption that it is only when an object looks that way under normal lightning conditions that it is red. Consequently I acquire a recognitional capacity for red things without ever having been causally influenced by one.

In any case, arguments about the practical difficulties in acquiring the requisite recognitional capacities carry little logical force regarding the nature of reference. They can all be refuted at once by considering again the swampman, who springs into existence with concepts which, like a brand new smoke alarm, stand in all the right informational relations without the swampman ever having been causally influenced by anything. I therefore conclude that when a concept is created it always stands in an informational relation to its reference but this does not require there to have been any causal interaction between the concept and the reference. Only the counterfactual causal relations needed to sustain the informational relation are essential.⁸

⁷ Dretske (1981, chapter 9) denied this, but only because he assumed that any concept introduced by a description would be a complex concept.

⁸ Fodor (1994, pp. 115-119) comes to fairly similar conclusions about some of the issues discussed in this appendix. He agrees that it is more plausible that the swampman shows that the causal history of a term is not relevant to its identity than that it is. Only for demonstrative thoughts, he argues, is the actual etiology of the term relevant to its identity; elsewhere it is the counterfactuals that count (p. 119).

APPENDIX B

Spatiotemporal Location and Reference to Physical Objects

I have suggested that Ways of Referring can sometimes be purely descriptive. The notion that spatiotemporal particulars such as physical objects can be identified using definite descriptions has, however, been criticised by Strawson (1959) and others, who claim that no spatiotemporal particular can be identified by description unless the description contains a demonstrative element (as in ‘the ball in *that* cupboard’). I disagree with this. Although I accept that no one could identify everything by description it is nonetheless possible in individual cases to identify a spatiotemporal particular using a descriptive Way of Referring without any supplementation by a demonstrative. The terms from which a definite description is constructed cannot themselves be descriptive *ad infinitum*, of course; such a description would have no content. But the terms of the description could pick out their referents via perceptual recognitional capacities; it is not essential for there to be a demonstrative component.

Strawson’s argument, which I shall assume is sufficiently familiar not to need stating in detail, turns on the idea that a description cannot be relied upon to pick out a unique individual because no matter how detailed the description is it is impossible to rule out the possibility of ‘massive reduplication’, where the description fits more than one item because of a duplication of a region of the universe. A demonstrative, by contrast, picks out a unique individual because demonstrative identification enables one to locate the particular within a spatiotemporal framework in which one is, oneself, situated. Although Strawson accepted that there could sometimes be ‘pure individuating descriptions’ such as ‘the first dog to be born at sea’ (1959, p. 26) which pick out a unique object by its non-spatiotemporal distinguishing features, he did not feel that reference to particulars could be achieved in this way. His reason for this seems to rest

largely on the fact that it would be impossible for someone to know that there was a unique object which satisfied the description. Moreover, even though it is logically possible that a definite description might happen to fit just one individual, thought which is achieved in this way would be of little use, according to Strawson. This is because it would be impossible to place the individual referred to within our general scheme of particular things, those which we can act upon and acquire new beliefs about. It would be impossible to learn anything new about an individual which is identified purely descriptively except by learning new general truths (1959, pp. 27-8).

I suspect that much of the intuitive appeal of Strawson's argument is due to the fact that in cases of massive reduplication, as well as in various similar examples derived from Kant, it is only the difference in the spatiotemporal locations of two objects that makes it possible to distinguish them. This line of thought led Evans (1982, p. 107) to suppose that spatiotemporal location is an essential part of what he called the 'fundamental ground of difference' for objects existing in space, that is, whatever distinguishes items of a given kind from other items. Consequently, Evans seems to have held that in order to identify an object at a given time it is necessary to be able to locate it in space and time. Demonstrative identification, of course, does this - by perceiving an object one is able to 'home in' on its location. It is sometimes also necessary to add a sortal, in order to distinguish different objects at the same spatiotemporal location such as a statue and a piece of clay, for example. These are different objects because one could be removed or destroyed without necessarily destroying the other.

Too much of the Strawsonian argument depends, however, on the observation that one can never be sure that there is just one item which satisfies a given definite description. For one can never be sure that just one item is picked out by a demonstrative either. It is true that demonstratives are much more reliable at capturing a unique reference than definite descriptions; one is much more likely to acquire true beliefs by identifying objects demonstratively and making perceptual judgments about

their properties than by constructing definite descriptions at random and guessing what should be predicated of them. But demonstrative identification is not infallible. Not only can a demonstrative fail to refer, there is also a possibility (at least in principle) of picking out more than one individual. There are at least two ways in which this could happen. Firstly, there could be an illusion in which the images of two different objects are superimposed. The objects would, of course, be at different locations in reality but they could at least appear to be at the same place. Secondly, there is no obvious reason why two objects could not, in principle, occupy the same location for a period of time. This may not happen in the actual world but it does not appear to be a logical impossibility. Objects could sometimes interpenetrate. This would make it impossible to regard spatiotemporal location as the fundamental ground of difference for physical objects; indeed, it would suggest that there is no such thing as *the* fundamental ground of difference. The objects would, of course, have to be differentiable in some way; perhaps in certain cases their velocities would suffice.

The possibility of interpenetration need not, incidentally, be in conflict with theories which suggest that the very idea of a physical object has built into it the idea of interaction through contact including, perhaps, contact with one's own body (cf. Peacocke 1993b). The interpenetration need not occur at every encounter; it would be coherent to imagine objects which had a certain probability of interpenetration, and a corresponding probability of resistance, for each contact.¹

Strawson's claim that the descriptive identification of individuals fails to integrate them into one's general world view in any useful way also seems no more than partially correct. Insofar as his aim was to show that it was impossible that the identification of particulars is *always* by description then he is no doubt correct. But surely if one

¹ The possibility of qualitatively identical objects existing at the same location has been argued for by Mark Johnston (1987) and apparently also in unpublished lectures by Saul Kripke. Peacocke (1992, pp. 234-5, 1993b, pp. 170-1) briefly makes a similar point to the one just made against Evans's insistence that there must be a fundamental ground of difference for all particulars, and that the fundamental ground of difference for physical objects is spatiotemporal location. Peacocke (1992, p.

identifies an object descriptively at a particular time there is nothing to prevent one from then recognising a certain perceived object as the one which fits the description. One is then able to learn more about the object or interact with it. The possibility of demonstrative identifications may be indispensable in order for there to be any point in identifying an object descriptively but this does not render the latter form of identification incoherent.

There is therefore no relevant difference in logical status between demonstrative and descriptive ways of identifying spatiotemporal particulars. There is nothing incoherent in the idea that some (though not all) concepts might have their references identified descriptively. Descriptive and demonstrative identifications simply differ in their degrees of reliability, and in the fact that the latter facilitate action and the relatively reliable acquisition of true beliefs.

235) also adds that there seems to be no fundamental ground of difference for times, and yet thought about them seems unproblematic.

CHAPTER 4

Concepts III: Communication

Up to now we have been considering concepts possessed by a single person. We should now consider whether the same concept can be possessed by more than one person. For advocates of Inferential Role Semantics this question poses relatively few problems; two terms express the same concept if they have the same inferential role regardless of whether the terms belong to the same person or to different people. For Conceptual Atomism the issue is a little more subtle. I shall argue that concepts can indeed be shared by different people but only when they engage in a communal word-using practice. Such a practice facilitates *transparent communication*, in which the mutual deference involved in the communal word-using practice creates rational relations between the thoughts of different people. These rational relations are of the kind which, according to the Transparency Principle, can only be explained by a sharing of concepts.

4.1 *Communication and shared concepts*

Frege's official view of communication was that it required the speaker and hearer to entertain the same thought. Since entertaining the same thought requires possessing the same concepts ('modes of presentation' in Frege's terminology) he thus held that communication required that the speaker and hearer expressed the same modes of presentation when they used the same words.

In his article 'On Sense and Reference' (1892/1966), however, in a footnote to the claim that 'the sense of a proper name is grasped by everybody who is sufficiently

familiar with the language or totality of designations to which it belongs', he seems to portray this view of communication as an ideal which is not always realised in practice:

In the case of an actual proper name such as 'Aristotle' opinions as to the sense may differ. It might, for instance, be taken to be the following: the pupil of Plato and teacher of Alexander the Great. Anybody who does this will attach another sense to the sentence 'Aristotle was born in Stagira' than will someone who takes as the sense of the name: the teacher of Alexander the Great who was born in Stagira. So long as the *Bedeutung* remains the same, such variations of sense may be tolerated, although they are to be avoided in the theoretical structure of a demonstrative science and ought not to occur in a perfect language (1892/1966, footnote 2).

As well as giving some of the strongest support for reading Frege as having advocated a 'descriptive' theory of sense, this passage seems to conflict with the idea that genuine communication always involves the sharing of thoughts. It offers a view of communication as requiring only that words are used with the same references (*Bedeutung*); the same thoughts might be shared but they do not have to be.

In some of his subsequent writings, however, Frege gave arguments for the sharing of thoughts in which it did not sound as though he found the idea that communication could involve differences in sense 'tolerable' at all. Both *Logic* (1997), which is thought to have been written in 1897, and *The Thought* (1956), which was published in 1918, contain arguments such as the following (from *Logic*):

A thought does not belong specially to the person who thinks it, as does an idea to the person who has it: everyone who grasps it encounters it in the same way, as the same thought. Otherwise two people would never attach the same thought to the same sentence, but each would have his own thought; and if, say, one person put $2 \cdot 2 = 4$ forward as true whilst another denied it, there would be no contradiction, because what was asserted by one would be quite different from what was rejected by the other. It would be quite impossible for the

assertions of different people to contradict one another, for a contradiction occurs only when it is the very same thought that one person is asserting to be true and another to be false. So a dispute about something would be futile. There would simply be no common ground to fight on; each thought would be enclosed in its own private world and a contradiction between the thoughts of different people would be like a war between ourselves and the inhabitants of Mars (1997, p. 235).

It is possible that Frege changed his views after he wrote 'On Sense and Reference'; he might have even changed his view about what sense is. What is of interest for our present purposes, however, is that the two passages quoted describe quite different views of communication. On the first view sharing the same thought is seen only as an ideal which is not essential for communication to take place; only a common reference is essential. The second view, however, contradicts this; it suggests that the very idea of communication requires both parties to entertain the same thought. Failure to do so results in them simply talking past one another.

Clearly anyone who advocates a descriptive theory of sense, or indeed any version of Inferential Role Semantics, must accept Frege's point in the earlier passage from 'On Sense and Reference' that different people may often associate different descriptions (or, more generally, different Ways of Referring), and therefore different senses, with a word. Advocates of these theories will hold, with the earlier Frege, that communication between people who associate different senses with a word occurs provided they refer to the same thing using that word.

A number of philosophers have indeed accepted this conclusion. Evans, for example, in his discussion of proper names, reasons as follows:

Is there any particular *way* in which one must think of the object?

It would appear not. Even if we refuse to acknowledge the great diversity in ways of thinking of the referent that are likely to be found among the consumers in a mature name-

using practice, there are undeniable differences between a producer, who knows the individual whom he can recognise as NN..., and a consumer, who does not.

When we contemplate the fact that different audiences think of the referent in a variety of different ways (which may be quite different from the way in which the speaker thinks of it), we do not necessarily have to conclude that the audiences do not know what the speaker is referring to.

It would appear, then, that the single main requirement for a use of a proper name is that one think of the referent (1982, pp. 399-400).¹

Elsewhere (e.g. 1982, p. 316) he says that we come closest to the Fregean ideal of communication when using certain (but not all) indexicals, though he still seems to cast doubt upon the idea that exactly the same thoughts are shared.²

For Inferential Role Semantics the problem with the Fregean ideal of communication is essentially an epistemological one; it relates to the fact that it can be difficult to tell whether someone else is thinking of the reference in the same way that one is thinking of it oneself. Nevertheless, as mentioned above, such theories make it perfectly clear what it means to say that two people do or do not possess the same concept. For Conceptual Atomism, however, there is a different problem; it is necessary to make it clear what it means to say that two people share the same concept.

Take Fodor's account, for example.³ According to Fodor coreferential concepts are distinguished from one another by syntax; that is, by being associated with physically different mental symbols. These symbols do not have to differ in any particular way, except that they have to differ in such a way as to produce the required cognitive

¹ See also Evans 1982, §1.9 for similar remarks. At the very end of chapter 11 of *The Varieties of Reference* (1982, p. 404) McDowell notes that Evans had made some notes towards a further discussion of this issue but was unable to reconstruct their drift with confidence.

² See also Dummett 1981 (p. 102f), who accepts that different speakers may employ different means to determine the reference and may therefore associate different senses with a word, but claims that this occurs only to a limited extent because 'only what is more or less common knowledge will normally be taken as part of that sense' (p. 103). Nevertheless this will still 'leave a great deal of play' (ibid).

differences between different concepts. Fodor denies, however, that this commits him to a version of type-physicalism, the identifying of psychological types with neurological types. Typehood can be defined functionally for a given system instead, such that different tokens of 'Mentalese' are of the same type (and would therefore instantiate the same concept) if they would have produced the same computations within a given system (Fodor 1994, pp. 105-109).

This, however, leaves open the question of what is it for two people to share the same concept. If the neurological tokens that instantiate the concept CICERO in my brain happen to be physically type-identical to the ones that instantiate the concept TULLY in your brain (where CICERO and TULLY corefer) does that imply that when I say 'Tully' I am expressing the concept that you express using 'Cicero', and vice versa? It is not at all obvious that this could not happen, given that other aspects of our neurophysiology might differ.

In *Concepts* (1998) Fodor made it a condition on an acceptable theory of concepts that 'concepts are *public*; they're the sorts of things that lots of people can, and do, *share*' (1998, p. 28). But it seems that at this point he had not noticed the problem discussed above because, in response to criticisms made by Peacocke (2000, p. 335), he now accepts that there may simply be no fact of the matter about whether a concept is shared:

What, if anything, makes A's thought that *Cicero is wet* type identical to B's thought that *Cicero is wet* rather than to B's thought that *Tully is wet*. (It clearly can't be the spelling, assuming the thoughts are in Mentalese.)... It's probably possible to construct cases in which any Direct Reference theorist (not just any Informational Atomist) has to say that there's no question of truth as to which of two necessarily coextensive, syntactically primitive names are translations (Fodor 2000, pp. 369-70).

³ See above, sections 2.4 and 3.1.

A corresponding question arises for any version of Conceptual Atomism. For the theory for which I have been arguing the question amounts to this: the theory individuates concepts in terms of referential episodes. But until now we have only been considering concepts possessed by a single individual and consequently it has only been necessary to consider a referential episode as something intrapersonal. Now, we could just leave it at that, much as Fodor does, and assume that referential episodes are inherently intrapersonal. The question of whether two people share the same concept could then only receive either a negative answer or perhaps no answer at all. The notion of a referential episode can, however, be extended beyond the intrapersonal case, and concepts can be shared. To see how this can be so we need to examine more closely the distinction we encountered in Frege's writings between two different kinds of communication.

4.2 *Interpretive versus transparent communication*

The first kind of communication, which I shall call *interpretive* communication, corresponds to the earlier of the two passages from Frege quoted above and is familiar from Davidson's (1984) notion of 'radical interpretation'. Radical interpretation involves trying to understand the utterances of a speaker of a foreign language by making hypotheses about the meanings of a speaker's sentences. If one were to write the hypotheses down one would typically use one's own language as a metalanguage. Hence Jones, when confronted by the alien Smith, might hypothesise that:

- (1) 'Roses are red' (in Smith's language) is true if and only if roses are red

Such hypotheses are collated in order to construct a Tarski-style truth theory which can then serve as a theory of meaning for Smith's language.

Tarski's theory makes truth a relation between language and the world rather than a relation between one language and another. Correspondingly, understanding Smith's language does not require *translating* it into one's own language but rather *interpreting* it; that is to say, determining the references of the words of Smith's language along with the way in which the words can be systematically combined to form sentences. It is, of course, clear that one must regard Davidsonian theories of meaning in this way if they are to provide an account of one's understanding of one's own language, for this understanding could not consist in the understanding of some other language.

Jones's theory of meaning for Smith's language must contain axioms giving the denotation of primitive terms in order to facilitate the recursive construction of theorems of the same general form as (1) for any sentence that Smith can produce. For example there must be axioms such as:

(2) The reference of 'Cicero' (in Smith's language) = 'Cicero'

In section 3.1 it was described how Evans and McDowell have shown that a theory of reference of this form can be used as a theory of sense. This requires choosing one particular formulation of the theory of reference, because there are many different interpretations which give equally correct theories of reference for a given language. The following clause, for example, would be just as valid as (2) in a theory of reference for Smith's language:

(3) The reference of 'Cicero' (in Smith's language) = Tully

It would not, however, capture the sense of Smith's word 'Cicero'.

Now, when Jones radically interprets Smith (i.e. treats Smith as an alien speaker) Jones makes hypotheses about the *references* of Smith's words. A theory of reference

is then gradually constructed for Smith's language. But it may be that this theory of reference captures only the references, and not the senses, of Smith's words. Provided Jones assigns the correct references to Smith's words, however, Smith and Jones can communicate in a certain way (Jones has to get the combinatorial structure right too, of course). It does not matter for this kind of communication that when Smith says that 'Cicero was a great orator', Smith and Jones think of Cicero under different modes of presentation; so long as they both think of *Cicero* that he was a great orator then in a certain sense they understand one another. This is what I call interpretive communication. Its key feature is that understanding a speaker's utterance always involves making a hypothesis about the reference of each of the speaker's words.

Let us now consider whether there is a way in which the Transparency Principle could be applied to interpretive communication to help clarify the relations between the concepts of two people. As a reminder, the Transparency Principle is as follows:

The Transparency Principle: If an attitude A to the thought expressed by the sentence $S(\gamma)$ can make it irrational to dissent from attitude A to the thought expressed by the sentence $S(\mu)$, without the involvement of any further thoughts, then ' γ ' and ' μ ' express the same mode of presentation.

Since rationality tends to be thought of as an intrapersonal rather than an interpersonal matter it may appear that this principle has no application to the interpersonal situation involved in communication. A version of the Transparency Principle for the interpersonal case would have to be something like the following:

Interpersonal Transparency Principle: If, when Smith takes an attitude A to the thought expressed by the sentence $S(\gamma)$, this can make it irrational for Jones to dissent from attitude A to the thought expressed by the sentence $S(\mu)$, without the

involvement of any further thoughts, then ' γ ' and ' μ ' express the same mode of presentation.

This is a little abstract. A more concrete example would be:

Example: Suppose Smith believes that 'Cicero was a great orator'. If there are circumstances under which Smith's believing this could make it irrational for Jones to believe that 'Cicero was not a great orator', without the involvement of any further thoughts, then Smith and Jones both express the same concept using the word 'Cicero'.

The clause 'without the involvement of any further thoughts' is to rule out cases where, for example, Jones believes something of the form: 'Smith's word "Cicero" refers to the same thing as my word "Cicero"'. Clearly if further thoughts of this kind can come into play then it will always be possible for there to be situations in which Smith's thought could make Jones's thought irrational; but this would not tell us anything about the individuation of concepts in those situations. This corresponds to standard intrapersonal cases like that in which my belief that 'Cicero was a great orator' makes it irrational for me to believe that 'Tully was a great orator', but only because I happen to believe that 'Cicero is Tully'. The rational relations between my thoughts would not, in those circumstances, imply that I expressed the same concept using 'Cicero' and 'Tully'.

The Transparency Principle therefore only has any relevance in circumstances in which one person's thought can affect whether or not another person's thought is rational. This does not occur in interpretive communication. Suppose that Jones makes the hypothesis that Smith's word 'Cicero' refers to Cicero. If Jones then believes that Smith's sentence 'Cicero was a great orator' expresses a truth it would be irrational for Jones to assert that 'Cicero was not a great orator'. But this rests on Jones's belief that

Smith's word 'Cicero' refers to Cicero. This is just an empirical hypothesis and there is no evidence which can make it rationally necessary for Jones to accept it. It is always rational, in other words, for Jones to entertain doubts as to whether or not Smith's word 'Cicero' refers to Cicero. Consequently Smith's thought that 'Cicero was a great orator' cannot make Jones's thought that 'Cicero was not a great orator' irrational.

We can draw an analogy with the dynamic case here. Suppose that earlier I believed that 'Cicero' was a great orator' and now I believe that 'Cicero" was a great orator'. If the only rational relation between the concepts expressed by 'Cicero'" and 'Cicero"' comes about because I make a hypothesis that my earlier utterance of 'Cicero'" referred to the same person as my current utterance of 'Cicero"' then we have no reason to conclude that both terms express the same concept. The rational relations implicit in the idea of a retained concept do not come about through a kind of radical interpretation of one's earlier utterances. On the contrary, these rational relations only come about when there is diachronic transparency between one's thoughts at different times. Interpretive communication likewise lacks this transparency so there is no reason to suppose that the different speakers share the same concepts.

The analogy suggests that different speakers could only share the same concepts if communication between them was transparent. Transparent communication would occur if Jones did not need to rely on the hypothesis that Smith's utterance of 'Cicero' referred to Cicero; it would correspond to the dynamic case in which the rational influence of my earlier thoughts about Cicero on my current ones does not rely on my making the hypothesis that my earlier utterance of 'Cicero' referred to Cicero. Instead the influence is direct and transparent.⁴ In fact transparent communication does take place but it depends on the existence of the communal word-using practices that make up a shared language.

⁴ An analogy could also be made with the way in which perception can be transparent when one takes one's perceptions at face value. One accepts the testimony of one's senses unreflectively, and in much the same way transparent communication can occur only when one accepts the testimony of a speaker unreflectively. See below, and also see above section 2.7.

4.3 *Word-using practices*

Davidson holds that communication is always interpretive; it is always radical interpretation:

It is an enormous convenience that many people speak in similar ways, and therefore can be interpreted in more or less the same way. But in principle communication does not demand that any two people speak the same language. What must be shared is the interpreter's and the speaker's understanding of the speaker's words (1986, p. 438).

He denies, in fact, that there is any such thing as a language insofar as knowing a language means knowing a set of conventions which both speaker and hearer must share in order to communicate. Each of us has a 'prior' theory based on the way in which we use words ourselves. We normally take the prior theory as an initial hypothesis about how other people will speak because this saves a lot of time, but this theory is soon adapted into a 'passing' theory specific to a conversation in order to take account of idiosyncratic and novel uses of language by our correspondent. Davidson argues that although idiosyncrasies such as malapropisms are relatively common they do not usually inhibit communication; when I say, standing before what is very obviously a swan, 'that duck looks hungry', it is very easy and natural to interpret me as using the word 'duck' to refer to a variety of waterfowl including swans. Idiosyncratic uses of language do not prevent us from communicating interpretively and this shows how little we need conventions in order to communicate. Understanding other speakers, for Davidson, is part of the more general task of understanding the world and it makes use of the same hypothesis generating abilities.

I accept that much of what Davidson says is correct; certainly we could not communicate without the ability to make hypotheses about the meanings which other speakers attach to their words. So there has to be such a thing as interpretive

communication, and it has to be pretty common. Sometimes, however, we communicate in a way which makes use of conventions even though these are not strictly necessary for the kind of communication with which Davidson is concerned. These conventions make communication easier and more efficient; and, moreover, they do so by making it transparent.

The conventions that I have in mind belong to what I shall call 'word-using practices'.⁵ It is of course no coincidence that the majority of people within a community use a certain word to refer to a certain thing most of the time. At some point in time a new word is created by someone. This word is then passed on to other people and it spreads through the community. Many of those who acquire the word acquire the means to identify its reference, but some do not, and must therefore defer to those who do. All of them take part in a practice of using a certain word to refer to a certain thing. So far this is, of course, a very familiar story.

Let us focus on the phenomenon of deference. As is well-known, many of us have to defer to 'experts' who are able to identify things which we are not able to; for instance those who, in Putnam's (1975) famous example, can tell the difference between elm trees and beech trees.⁶ I am, in fact, among that group of people who cannot distinguish elms from beeches but I have no idea who the relevant 'experts' are. Fortunately, I do not have to know this explicitly; provided I can identify the community of users of the word 'elm' then it will not be too difficult to find someone

⁵ Evans (1982, chapter 11) describes 'name-using practices'; my idea is similar (though not in every detail) and I use a different terminology only to emphasise that my account is intended to apply to all categories of words whose reference does not depend on context. In fact, I am oversimplifying a little here for expository purposes. It will become apparent that a word-using practice is a special case of a *referential practice*, the individuation of which is not quite so closely tied to the use of a particular word. This allows, for example, an arbitrary change in the word used to refer to a particular thing without a change of concept (e.g. 'starting from tomorrow we shall use the word 'green' to refer to red things and vice versa...'). We shall need to appeal to the more general notion of a referential practice when we come to consider communication using indexicals, where one word is systematically exchanged for another when there is a change of context.

⁶ See also Burge 1979. Burge's examples show that the reference of a term can be determined by what 'expert' members of a linguistic community use it to refer to even when some members of the community are under a misapprehension about the reference.

in that community who knows how to identify elms. I do not have to know who that person is in advance.

So deference is not so much a matter of knowing who the experts are as knowing which community they belong to. Consequently participation in a word-using practice requires deference to the community whose use of the word constitutes the practice. Now, deference is often portrayed as one-way traffic; those who cannot single-handedly identify the reference have to defer to those who can but the latter, the experts, do not defer to anyone.⁷ This is not quite correct; there are generally no experts whose use of a word is entirely independent of other people's use of it.

Consider my use of the word 'dolphin', for example. I can, as a matter of fact, identify dolphins. So, strictly speaking, I am an expert. But I am not sure that I would like to claim that my use of the word 'dolphin' cannot be mistaken. I simply believe that I know how to identify dolphins. For all I know, it might turn out that I do not know how to identify dolphins at all, and the creatures that I thought were dolphins are actually sharks. So even supposing for the sake of argument that I do as a matter of fact identify dolphins correctly I still have to be prepared in principle to defer to others.

This would be true even if I had introduced the word 'dolphin' into the language myself. Admittedly, at the moment at which I first coined the term my use would be definitive and I could not be mistaken; I would be stipulating that 'dolphin' refers to *that* species of marine animal, and stipulations cannot be false. But it is, of course, possible that after some period of time my use of the term might change without my realising; due to a poor memory I might start calling sharks 'dolphins', for example.

If no one else had been introduced to my word 'dolphin' then if I changed my use of the term and called sharks 'dolphins' we should simply say that I had come to express a different concept.⁸ But suppose that I had introduced other people to the 'dolphin'-

⁷ Cf. Evans's (1982, chapter 11) distinction between 'consumers' and 'producers'.

⁸ Some philosophers, particularly those influenced by Kripke's (1982) interpretation of Wittgenstein (1958), have used considerations such as the fact that an individual's reference can drift to argue that the required normativity of word-use is only possible within a linguistic community. But, quite apart

practice before my use changed. If none of them had learned to identify dolphins except by deferring to me then when my use changed so would theirs. But this would be an unusual situation; it would be more common for at least some people, on learning the word 'dolphin' from me, to simultaneously learn how to identify dolphins. Teaching them this would typically, though perhaps not necessarily, involve getting them to refer to dolphins using a Way of Referring which I used myself. For example if I pointed to a dolphin and said 'that species of marine animal is called a "dolphin"' the learner might acquire a recognitional capacity for dolphins comparable with my own. Or if I told them that 'dolphins are distinguished by such-and-such features' then they would share with me a descriptive Way of Referring.⁹ It is probably not crucial that they share the same Way of Referring with me, however, so long as they learn how to identify the reference correctly.

Even if we started with the same Ways of Referring, however, our Ways of Referring could subsequently diverge. This would not be not a problem provided we continued to refer to the same thing. Now, if at some point I started misidentifying dolphins then the other members of the linguistic community could correct me on this. In order for this to be possible it would be necessary for me to be willing to defer to them; I would have to accept that I no longer had a monopoly on the use of the word 'dolphin' and that it was now public property. I could, of course, choose to ignore other people. But this would be a case in which other people borrowed a term from me in starting a word-using practice in which I did not participate.

As a practice continues it becomes unnecessary to remember who its originator was. The practice can continue for as long as the members of the linguistic community continue to refer to the same object or property. Individuals may go astray by using the

from the fact that the use of a word by a whole community could in principle drift, normativity is surely accounted for simply by the fact that when a word is used with a given reference there is a particular item relevant to the truth or falsity of what is said; the possibility of unnoticed drift would not threaten the truth or falsity of particular uses of a term.

⁹ This is to some extent comparable with the temporal 'overlap' in Ways of Referring required to retain a concept through time. See above, section 3.2.

word with the wrong reference, but can be brought back into the practice by being corrected by others. Rather than consisting of a set of experts who defer to no one and a set of 'consumers' who defer to the experts, then, a word-using practice is constituted by a community of users of the word all of whom are willing in principle to defer to other members of the community.¹⁰

Let us return to the case of Jones trying to understand Smith. In order for it to be possible to recognise a structure in Smith's utterances, and in order to be able to recognise the same elements of the structure (i.e. the same words or phrases) when they reoccur in other utterances or in other parts of the same utterance it is necessary for Jones to be able to identify Smith's words. This is not always a trivial matter; many of us have had the experience of listening to a foreign language being spoken and not being able to identify individual words. This can even happen to some extent with a familiar language when it is spoken with an unfamiliar or foreign accent. Having made a hypothesis about an initial utterance one then has to make a hypothesis for each subsequent utterance about whether the same words have been uttered. So it is always necessary to make a hypothesis about which words have been uttered. Furthermore, it will not do just to be able to identify generic words.¹¹ Smith might use the same generic word to refer to two different things, so it is necessary to be able to determine

¹⁰ Fodor (e.g. 1994, p. 33; 1998, p. 154) holds the asymmetric view of deference against which I am arguing. He could, however, adopt my story without having to renounce anything important; for what matters to Fodor is that someone in the community must be able to identify the reference independently, otherwise the whole thing goes round in a circle. My point is that being able to identify the reference independently does not exempt one from deferring to others, if 'deference' is understood in a certain way. Woodfield (2000) argues against Fodor's asymmetric view but for different reasons to those given above. In fact, as Recanati (2000) points out (partly in response to Woodfield), it is important to distinguish different kinds of deference. Woodfield's remarks mainly concern cases in which someone possesses a concept independently of the experts but still defers to them in difficult cases. If, for instance, you have better eyesight than me then I may defer to you about whether an individual seen in the distance is Jones even though I might be able to identify Jones from close up. I am not convinced that this kind of deference has anything to do with the individuation of concepts. In any case, the kind of deference with which I am concerned (and presumably this applies to Fodor as well) is that in which it is at least possible (for all I know) that I cannot identify Jones at all, and my only access to Jones is through those people who can.

¹¹ On Kaplan's notion of generic words see above section 3.1.

which *use* of the generic word (or which Kaplanian common currency ‘word’) is involved in the utterance. This will normally be determined by context.

If Smith and Jones are communicating interpretively there will therefore be two steps involved in understanding utterances: Jones has firstly to identify the words which Smith uses, and then make a further hypothesis about what each of the words refers to. An error could occur in either of these steps. Having made an initial hypothesis about the reference of a word Jones may then assume that subsequent occurrences of that word have the same reference, but Jones’s ascription of meaning to each utterance still rests on the initial hypothesis.

Suppose, however, that Smith and Jones both belong to the same linguistic community and exchange only words which belong to the word-using practices of that community. Now their communication becomes transparent rather than interpretive. It is still necessary for Jones to make a hypothesis about which word Smith is uttering but, unlike in interpretive communication, no further hypothesis needs to be made. When Jones determines that Smith has uttered the word ‘Cicero’ - a word which already belongs to Jones’s vocabulary - there is nothing further to be determined because Jones already knows what ‘Cicero’ refers to.

This is clearest in a case in which Jones defers totally to Smith about the identity of Cicero; a case in which Jones’s only way of identifying Cicero is through Smith. In that case it would make no sense, upon hearing Smith’s utterance of ‘Cicero was a great orator’, for Jones to think ‘Smith is saying that someone is a great orator, but is it *Cicero* about whom Smith is saying this?’¹² Consequently, if Jones were to take it for granted that Smith was speaking the truth it would be irrational for Jones to dissent from the thought that ‘Cicero was a great orator’. Thus, by appeal to the Transparency Principle as described above, we can conclude that Jones must share Smith’s concept

¹² The case is structurally similar to a form of ‘immunity to error through misidentification’ (see section 6.6) which occurs with demonstratives: it makes no sense to think, based on perceptual contact with an object, ‘something is *F*’ but is it *that* which is *F*?’ (where the demonstrative identification makes use of the same channel of information upon which the predication is based).

CICERO. This is the only way to account for the rational relations between Smith and Jones's thoughts.¹³

Since deference is not limited to those who lack an independent means by which to identify the reference, however, transparent communication can occur between any members of a word-using practice, whether or not the correspondents rely on experts to identify the references of their words.¹⁴ Even if Smith and Jones both have a way of identifying Cicero, and both do so correctly, Jones still needs no hypothesis about the reference of Smith's word 'Cicero'. Jones simply has to correctly identify the word-using practice to which Smith's utterance of 'Cicero' belongs. Transparent communication thus corresponds to the Fregean ideal of communication in which both speaker and hearer entertain the same thought; this allows disputants to directly contradict one another, as described in the later quotation from Frege, above.¹⁵

Let us briefly recap. I have argued that word-using practices involve mutual deference by all members of the practice. Because of this, there are rational relations between the thoughts of different thinkers, and these are of precisely the kind which

¹³ Similar intuitions may lie behind Richard G. Heck Jr.'s (1995) argument for the sharing of senses. Heck's argument is complex but the main thrust of it is that communication is essentially a means for the transmission of knowledge, and mere coreference does not suffice for this. If the acquisition of knowledge is assumed to require reasons then Heck is essentially arguing that communication must be transparent in much the way that I have been describing.

¹⁴ Consequently it does not matter for our present purposes whether or not we say that someone who defers entirely to others really possesses the concept. I confess that I am uncertain of the correct treatment of this issue. It would be possible to argue that there are circumstances in which someone can, by virtue of deference to others, use a word without understanding it or associating any concept with it (cf. Evans 1982, pp. 400f). Certainly if the only thing Jones believes about Cicero is that Cicero is whoever Smith (and other members of the 'Cicero'-practice) refers to as 'Cicero', then it is debatable whether Jones can really be said to have the means to satisfy Russell's Principle. On the other hand I have a strong intuition that I have the concepts ELM and BEECH even though I cannot tell elms and beeches apart. The question is what is involved in being able to identify something. I can presumably still identify something if I need to put glasses on to do so. So tools of certain kinds can be used. But what if I have to carry a book on trees with me in order to distinguish elms from beeches? And if that counts, why not think of other people as tools which I can use to help me identify things? I am not sure how to answer these questions in a principled way, and I leave the matter for further research.

¹⁵ Since in order to defer it is essential to know who one defers to my view contrasts with Kripke's (1980, p. 92) 'causal' view, in which what matters is who Jones learned the word from (learning is assumed to involve causation) even if this is not who Jones believes that the word was learned from. On my view the swampman (see Appendix A) can defer to a particular practice and therefore share a concept even without having received any causal influence from it.

the Transparency Principle shows must be explained by shared concepts. Consequently, different speakers can share concepts even though they may differ in the Ways of Referring that they associate with a given concept.¹⁶ The notion of a referential episode can therefore be extended to the interpersonal case; as new members join a word-using practice the referential episode spreads across the linguistic community. As with the dynamic case, incidentally, it is important not to misunderstand the structure of the argument with respect to the Transparency Principle. As emphasised above, the Transparency Principle sets constraints on the individuation of concepts but it does not constitute a theory of concepts itself.¹⁷ In this particular case, that Transparency Principle shows that if there are rational relations of a certain kind between the thoughts of different thinkers then those thoughts must be the same. But it is the account of the extension of the notion of a referential episode through word-using practices which yields the individuation of concepts which gives rise to the rational relations in question.

A number of situations can occur which may appear to give rise to counterexamples to this account. For example it may appear that it would be possible for Jones to defer to Smith without Smith deferring to Jones. This would be a problem because it would imply that Jones shares Smith's concept but Smith does not share Jones's concept. Fortunately, this situation cannot occur. In order for it to be appropriate for Jones to defer to Smith's use of 'Cicero', Jones must have acquired the word from Smith.¹⁸ Smith must therefore have used the word in communicating with Jones. But once Smith has communicated the word to Jones (and assuming that Jones may have

¹⁶ It would in fact have been possible to start by insisting, as Frege does in *Logic* (in the quotation above), that transparent communication clearly does take place and then draw from this the conclusion that sameness of concept cannot require sameness of Way of Referring. The structure of the argument would then have mirrored the structure of the argument given for the dynamic case in chapter 2. I am not, however, entirely sanguine about the chances of convincing anyone by putting the argument that way around.

¹⁷ See section 1.2.

¹⁸ Normally Jones would of course defer to a whole community of word-users, of whom Smith is just one member. I am discussing a community consisting of just two speakers in order to keep things simple, but the conclusions can be straightforwardly generalised for larger communities.

acquired the means to identify the reference), Smith is rationally obliged to defer to Jones if Jones uses the word. For, having learned the word, Jones would be in a position to correct Smith if Smith's use went astray. In allowing others access to one's terms, which one does whenever one uses them in public, one opens oneself to the possibility of being corrected in future uses.

What if Smith is a hermit who never talks to anyone and never even acknowledges their existence? Perhaps Jones overhears Smith mumble something about 'Cicero' and decides to defer to Smith about its meaning. But this is not really deference; Jones does not *learn* the word from Smith. This is just a case in which Jones introduces a term as referring to 'whoever Smith refers to using "Cicero"'. The most Jones can do in that situation is to interpret Smith. One cannot defer to someone unless they are willing to correct one's own use of the term. A word-using practice is like a game which requires all of the players to participate; it requires mutual co-operation. One cannot be playing the game if one is in a position to change the rules unilaterally (e.g. by deciding to change the reference of the word), and in such a case one cannot be played with by others either.¹⁹

Another objection runs as follows. If Jones defers to Smith then Jones must believe that Cicero is whoever Smith refers to as 'Cicero'. Why can we not regard this as Jones's Way of Referring to Cicero?²⁰ And if it is a Way of Referring then according to the theory proposed in the previous chapter it should be possible to exchange it for a different Way of Referring without a change of concept. So suppose Jones discovers that Cicero is the individual who is uniquely ϕ . Jones may then forget all about Smith's use of 'Cicero' but continue to refer to Cicero in virtue of Cicero being the ϕ . But the rational relations between Smith's thoughts and Jones's would then disappear, which

¹⁹ A rough analogy in the intrapersonal case would be this: suppose I believe that everything that is true of *a* is also true of *b* but I do not believe that everything that is true of *b* is also true of *a*. Then '*a*' and '*b*' do not express the same concepts. Sometimes the rational relations between two thoughts have to cut both ways in order for the thoughts to contain the same concepts. Cf. Frege's notion of 'equipollence' as a criterion for sameness of thoughts, section 1.1 above.

would imply that they were no longer the same thought. Worse still, it appears that Jones's belief that 'Cicero is whoever Smith refers to as "Cicero"' could explain the putative rational relations between Smith's and Jones's thoughts. If Jones believed that Smith's utterance of 'Cicero was a great orator' expressed a truth and Jones also believed that Cicero was whoever Smith referred to as 'Cicero' then Jones could *infer* that Cicero was a great orator. This is the kind of 'further thought' that has to be ruled out for the Transparency Principle to be applicable. Consequently, according to this objection, there is no need to say that Smith and Jones possess the same concept.

This objection implies that what I have been calling transparent communication is really just interpretive communication, and that the putative transparency between the thoughts of Smith and Jones can be explained away in terms of what Jones has to believe in order to defer to Smith. Now, it is certainly true that interpretive communication can lead one to make inferences of the sort described. To make this a little clearer it may help to imagine that when Smith uses the name 'Cicero' Jones interprets this as referring to Tully. Given that Jones believes that Smith's word 'Cicero' refers to Tully, when Jones believes Smith's statement that 'Cicero was a great orator' Jones can infer that Tully was a great orator. But that does not suffice to make Smith's 'Cicero' and Jones's 'Tully' express the same concept; the two concepts do not stand in a transparent relation to one another. By the same token, the objection argues, neither do Smith's 'Cicero' and Jones's 'Cicero'.

The very fact that we can exchange 'Cicero' for 'Tully' like this ought, however, to raise suspicions, for Jones could hardly defer to Smith's use of 'Cicero' while expressing the concept thus acquired using the word 'Tully'. The objection is mistaken but the reason for this is a subtle one. The best way to illustrate this is by returning to the analogy with the dynamic case. The retention of a concept through time can be seen as involving something analogous to deference with respect to one's earlier use of

²⁰ Though see footnote 14, above, on an uncertainty over whether such a 'Way of Referring' would suffice for concept possession.

the concept. In thinking about Cicero I take it for granted that I am now referring to the same individual that I was earlier. This is implicit in the fact that if, earlier, I thought that Cicero was a great orator then, all else being equal, I should still think this now (or at any rate I should not deny it).

It would be quite incorrect to explain this quasi-deferential relation by saying that I must normally believe something of the form 'the person I now refer to as "Cicero" is the same one that I referred to earlier as "Cicero"', and that my continuing to believe that Cicero was a great orator rests on an inference of the form: 'Earlier, I truly believed that "Cicero" was a great orator; by "Cicero" I was referring to Cicero; therefore Cicero was a great orator'. The inference is valid, but in the normal case it is entirely superfluous. I can simply retain the concept *CICERO* and continue to believe the same things about Cicero non-inferentially. The mere possibility of making such an inference does not show that such inferences are always necessary. By analogy, when Smith and Jones both belong to a word-using practice it is possible for Jones to hear Smith say that 'Cicero was a great orator' and thereby entertain the thought that 'Cicero was a great orator' non-inferentially. Understanding in this situation is direct and non-inferential; to identify Smith's sentence is to entertain Smith's thought. So the objection is not valid.²¹

Finally, I must make one small point about word-using practices. As with so many epistemological matters, there are no strict rules about how to identify an uttered word as belonging to a particular word-using practice. There might be more than one 'Cicero', in which case it would be necessary to use contextual clues to determine which one was being spoken about. As mentioned above, there may also be idiosyncratic pronunciations which have to be understood. There will always be an element of guesswork; most of the time we guess correctly though errors do, of

²¹ The objection just discussed is analogous to what McDowell has, in various publications, called the 'highest common factor conception'. It assumes, incorrectly, that just because interpretation can sometimes get in the way of transparent understanding it must always get in the way. See for example McDowell 1982a; see also McDowell 1993.

course, occur. It may be that malapropisms can sometimes be allowed for; if Smith accidentally says 'Sisley was a great orator' it may be possible for Jones to correctly judge that Smith meant to say 'Cicero'. Davidson may be right that there are no inflexible conventions with regard to orthographic words, but there can still be a uniform practice provided speakers can identify the practice.²²

A reasonably uniform pronunciation of a word is therefore important in enabling different speakers to share a practice, but the uniformity should not be exaggerated.²³ It is the practice, rather than a particular sound or symbol, which must be identified. Consequently it may be beneficial to re-describe what I have been calling a 'word-using practice' as a *referential practice*. This is a referential episode which has spread to more than one person. Just as an intrapersonal referential episode is not strictly tied to the use of a particular word, neither is a referential practice. In particular, when two speakers communicate using indexicals there may be changes of context which necessitate changes of indexical, but so long as both speakers understand the rules of use for indexical terms it is possible for a speaker who hears a series of indexical terms to realise that what is being expressed is part of a continuous referential practice.²⁴

²² Although we might be able to make sense of Smith saying 'Sisley' instead of 'Cicero' Smith must still defer to other people's uses of 'Cicero'. If Smith defers to utterances of 'Sisley' then it does not seem possible to regard Smith as a member of the 'Cicero'-practice.

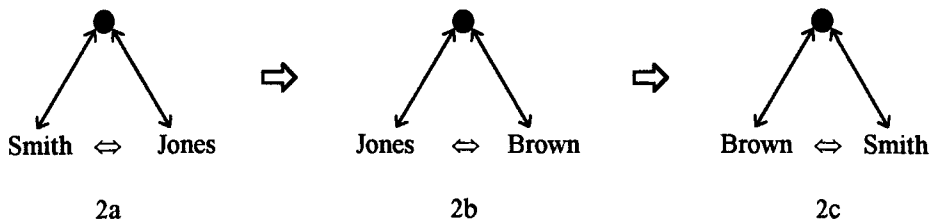
²³ When thinking becomes a communal exercise there may be rather more breakdowns in the rational integrity of thought than occur within a single person at a single time. Suppose, through deference, A shares a concept with B and B shares the same concept with C. Then A and C should share that concept. In that case, A and C *should* defer to one another's use of the corresponding word. But they might not actually do so; C may pronounce the word in a way which B can recognise but A cannot, and vice versa. Rather than say that A and C do not share the concept I think we should just say that they fail to recognise the rational relations in which their thoughts stand. It is partly because of this sort of thing that interpretive communication is important.

²⁴ This will be discussed in greater depth in chapter 5.

4.4 Demonstratives and the Paderewski paradox

The principles just described can be applied to communication using demonstratives. Smith and Jones can of course use demonstratives independently of one another in referring to the same object, and in this case they would possess different concepts. But if Smith says ‘that is blue’ Jones may respond anaphorically by saying ‘it is not blue’. The anaphoric ‘it’ can only refer to the same thing that Smith referred to using ‘that’. The notion of anaphor is akin to that of deference; anaphoric communication is transparent. If, instead, Jones had independently identified the object and said to Smith ‘that is not blue’ there would be no disagreement between them unless Jones intended the word ‘that’ to be understood as anaphoric on Smith’s demonstrative (the pragmatics of communication can presumably allow this in certain contexts). Smith can then respond back to Jones anaphorically or by using a further demonstrative which, once again, must be understood as anaphoric on earlier uses. A referential practice is thus initiated, in which Smith and Jones share a concept.

Demonstratives give rise to a puzzle which is in fact a special case of a phenomenon which applies to words of all kinds and stands in need of an explanation. The puzzle occurs when a participant in a referential practice fails to recognise another speaker’s participation in the same practice and assumes that it is a new practice. This is diagrammed for the case of demonstratives in figure 2.



In figure 2a Smith and Jones are jointly attending to an object and discussing it using demonstratives and anaphors. Smith and Jones mutually defer and thus share a concept which refers to the object. In figure 2b Smith has left the scene and no longer possesses a demonstrative Way of Referring to the object. But as a matter of fact Smith retains an alternative Way of Referring to the object (either descriptive or recognitional) and, moreover, would still defer to Jones if the occasion arose. Consequently Smith still retains the concept.

Another person, Brown, now appears and joins the referential practice with Jones. So Brown now shares the same concept with both Jones and Smith. At a still later time, depicted in figure 2c, Jones has also left the scene but Brown has continued to keep track of the object. Brown can continue to retain the same concept. In effect, Smith, Jones and Brown take it in turns to keep track of the object. But now Smith returns to the scene but does not recognise the object as the one encountered earlier. Consequently when Brown talks about 'that' Smith does not realise that Brown is talking about the object seen earlier. But Smith can communicate with Brown about the apparently new object and, on the face of it, Smith ought to be able to acquire the concept which Brown possesses. But this concept was supposed to be the one which Smith already possessed, having retained it from the earlier encounter. Yet it cannot be the concept which Smith now shares with Brown, because for Smith there is an informative identity between them.

The paradox just described has the same form as a paradox described by Saul Kripke in 'A Puzzle About Belief' (1979). I shall call it the 'Paderewski paradox'. Paderewski was celebrated both as a musician and as a politician. Suppose that Jones hears about Paderewski as a musician. Jones thus acquires a certain concept which involves deference to the linguistic community who use the name 'Paderewski'. But now suppose that in a different context Jones hears about Paderewski as a politician. Jones does not assume that the politician called 'Paderewski' is the same person as the musician called 'Paderewski'. So for Jones the identity:

(4) Paderewski (musician) = Paderewski (politician)

is informative. Jones therefore possesses two different concepts. Yet both occurrences of 'Paderewski' ought to express the same concept, as there is only one name, and only one concept, in general circulation in the linguistic community. The structure of the puzzle is much the same as the one described above for demonstratives because in both cases someone thinks that there are two referential practices where in fact there is only one.²⁵

If we are to say that just one of Jones's 'Paderewski' concepts is shared with the community then we need a principled way of deciding which one it is. The situation looks puzzling precisely because both concepts seem to have an equal claim to being shared. But there is one asymmetry between them. When Jones learns of Paderewski as a musician the situation is absolutely normal; if anything counts as a case of sharing a concept with the community, this should. But when Jones learns of Paderewski as a politician the situation is not the same; Jones approaches the conversation from which the concept is acquired already possessing the concept normally expressed by 'Paderewski'. When Jones assumes that Paderewski the politician is not Paderewski the musician this implies that Jones believes in a non-existent referential practice and tries to defer to it. It is impossible to participate in the one and only 'Paderewski' referential practice while consciously thinking that the practice in question is a different one from the Paderewski (qua musician) practice. So we must conclude that Jones does share one concept with the community - the first concept acquired - and that any

²⁵ The Paderewski paradox is similar to the famous paradox concerning Pierre's beliefs about 'London' and 'Londres' from the same article. It is possible, in my opinion, to solve the Pierre paradox by assuming that 'London' and 'Londres' express different concepts (and are involved in different referential practices, though admittedly these practices probably share common historical roots); that is why Pierre can believe that 'Londres est jolie' and 'London is not pretty' without contradiction. Many people seem to find it intuitively compelling that 'London' and 'Londres' must express the same concept, but I should have thought that Kripke's puzzle about Pierre makes it obvious that they do not. Gerald Vision (2001, p. 8), for example, in the process of criticising Fodor, takes it as a 'presumptive condition about concept possession' that monolingual speakers of different

further concepts are Jones's own creations and are not shared with anyone. This is true even though Jones may have conversations with other people about 'Paderewski' while thinking that they are talking about the politician, not the musician. Communication may take place at the level of reference but there is only an illusion of transparency and of the sharing of concepts.²⁶

Much the same can be said about the case of demonstratives discussed above. When Smith returns to the scene Smith excludes the possibility that Brown's referential practice is the same one which Smith had started earlier and regards it as a new one instead. Consequently Smith tries to enter a non-existent referential practice. Smith's new concept is therefore a brand new creation and is not shared with Brown.

languages can share concepts. But, given that concepts are thought components (which is assumed by Vision, as well as here) this assumption appears quite unjustified.

²⁶ This is not the only case in which there can be an illusion of transparency. Another one was discussed in section 1.1, in which an object being kept track of is switched for another one without the subject noticing. It was observed that this cannot be seen simply as a switching of concepts because a concept expressed using a demonstrative after the switch cannot be seen as independent of the concept expressed earlier.

Part 3: Indexicals

CHAPTER 5

Changing Indexicals

In a much discussed passage, Frege claimed that the same thought can sometimes be retained when one indexical term is exchanged for another to compensate for a change of context:

If someone wants to say the same today as he expressed yesterday using the word 'today', he must replace this word with 'yesterday'. Although the thought is the same, its verbal expression must be different in order that the change of sense which would otherwise be affected by the differing times of utterance may be cancelled out (Frege 1956, p. 296).

Gareth Evans (1981, 1982) dubbed these retained thoughts 'dynamic thoughts' and he, and subsequently several other philosophers, defended their existence.¹ For reasons that will become apparent in the next chapter, although Frege and Evans expressed their point in terms of retained thoughts I shall defend only the more modest claim that the same singular mode of presentation can be retained and expressed using one indexical term followed by another, and hence that a referential practice (and therefore a concept) can cut across different contexts and different indexical terms. This claim follows from the same principles of transparency that were used throughout Part 2 but since it may seem counterintuitive - especially to advocates of Inferential Role Semantics - I shall defend it in detail, examining several different arguments in turn. A serious objection, based on the fact that different indexical terms are naturally thought of as paradigm examples of terms with different inferential roles, is then dealt with in

¹ For further arguments see for instance Hoerl 1997, Luntley 1997, 1998, 1999.

the next chapter. This leads to the development of a novel theory of indexical thought which is compatible with Conceptual Atomism.

5.1 *Evans's argument*

Evans offers two arguments in favour of Frege's claim. The more prominent of these arguments is intriguing and demands discussion though despite agreeing with its conclusion I do not think it is valid as it stands. Evans follows Frege in focusing on an example involving temporal indexicals but both authors make it clear that their arguments apply equally to spatial indexicals such as 'here' and 'there'.² Since it is a little clearer I shall concentrate on the latter case. Suppose a subject 'S' has a thought of the form:

(1) It is *G* here³

S then moves away from the place while keeping track of it (typically using an ongoing perception of spatially stable objects or features). From a different place, S then thinks:

(2) It is *G* there

The word 'there' in (2) refers to the same place as the word 'here' in (1). Evans contrasts the Fregean claim that sentences like (1) and (2) represent 'cross-sections of a persisting belief state' with what he calls the 'atomistic' conception (which is not to be confused with Conceptual Atomism) under which (1) and (2) represent a series of

² See Frege (1956, p. 296), Evans (1985, pp. 310-311).

³ Grammatically many indexicals, including 'here' and 'now', are adverbs. I assume that this is a quirk of English which disguises their logical form, though perhaps an interesting one if they involve a kind of disguised predication as the theory which I shall advocate in chapter 6 suggests.

related but independent atomic thoughts. According to Evans on the atomistic conception 'the atom must be a perfectly coherent unit of thought by itself, even if it is entertained by one who has not the least propensity to form the other members of the series' (1985, p. 309). He then observes that this cannot be correct. For suppose that S were to think, in a certain place, that 'it is *G* here'. If S then moved a short distance to the left, say, it would then be normal for S to believe that 'it is *G* to the right'. In fact if someone had no propensity ever to do this then we should not credit them with thoughts about space at all, for if someone properly understands the notion of space they must surely be committed to the existence of a system of relations such that if one moves to one's left then the very same place where one was before is now to one's right. This is closely related to the claim that Peacocke (1997, p. 243) makes in arguing that PLACE and SPATIAL RELATION form a 'local holism' (see above, section 3.3). By the same token, to use Evans's example, if someone thinks 'it is now ψ ' then we cannot credit them with genuine temporal thought unless they have a propensity to think, a moment later, that 'it was ψ just a moment ago'; someone who could think the first thought without any propensity to think the second just wouldn't be thinking about a time. The propensity to form these thoughts in a series as the context changes is therefore a condition for the possibility of thought about space and time. We could say, borrowing some more of Peacocke's (1986) terminology, that it is among the canonical commitments of thoughts about places and times that the thinker is disposed to form a series of thoughts in this way.

This is consistent with what Evans says in section IV of his paper 'Things Without the Mind' (1980). Someone who could have one of the supposed atomic thoughts without a propensity to form the next one would have only what Evans calls a 'serial' grasp of space or time because they would not be committed to the mind-independent existence of places and times. By contrast Evans argues that a 'simultaneous' conception is essential for spatial thought (and likewise presumably temporal thought).

Analogous considerations can be applied to thought about physical objects; one cannot really be thinking about an object if one never expects its existence to persist. This notion of persistence lies behind the idea of keeping track of a perceived object - one would not really be keeping track in the relevant sense if one believed that what one was perceiving at each moment was an entirely independent time-slice rather than a temporal stage of an object which persists from one moment to the next. Thus one cannot fully manifest in an instant the fact that one is thinking about an object via a perceptual identification. This may lie behind Evans's claim that 'we have to regard the static notion of "having hold of an object at t " as essentially an abstraction from the dynamic notion of "keeping track of an object from t to t' "' (1985, p. 311). Much the same would apply to places and times.

Evans concludes that the supposed atomic thoughts cannot be atomic at all because they are not coherent independently of thoughts at other places and/or times. Hence, he concludes, the putative series of thoughts must instead be a single dynamic thought. There are two remarks I would like to make about this argument. Firstly, it is not absolutely clear from what has been said so far that the series of thoughts described by Evans must all involve precise reference to the same place or time. When someone thinks that 'it is G to the right' because of an earlier thought that 'it is G here' it is not obvious that it is a requirement that they must be thinking, of the exact same place, that it is G rather than simply thinking that it is G at a place that is now located somewhere to the right - a place which the subject may no longer be able to locate with precision.

One could perhaps try to supplement Evans's argument by saying that at least insofar as one keeps track of a place while moving around one must have a propensity to think the next thought in the series, for example to move from thinking (1) to thinking (2). My second remark, however, is that even with this supplementation Evans's argument is not valid. Evans shows that there are certain thoughts which are not possible unless the thinker has a propensity to think certain other thoughts, but

from what has been said it is not obvious why these subsequent thoughts must be identical to the earlier ones. There are many cases in which the ability to think a certain thought implies a propensity to think other, different thoughts. No one can be credited with the thought that 'there goes a rabbit' unless they are inclined to judge that 'rabbits are physical objects', yet the latter thought is obviously not identical to the former. Admittedly these thoughts do not form any kind of sequence, but Evans offers no argument as to the relevance of thoughts occurring in a sequence, or in a way which depends systematically on changes of context. Put abstractly, Evans's argument has the form: 'anyone who thinks "*a* is *F*" must have a propensity, when the context changes in a certain way, to think "*b* is *F*"; therefore '*a*' and '*b*' express the same mode of presentation'. Clearly without further supplementation this is not valid.

5.2 Transparency

It may be, however, that what Evans was really getting at in the argument just discussed was the fact that, as he later put it:

It is a precondition of rationality that information acquired at one time should be available to the subject later: hence, given that in a rational creature information (and misinformation) generates beliefs, that beliefs should persist (1982, p. 235).

This hints at the notion, embodied in the Transparency Principle, that there are rational relations between thoughts at different times which can only be accounted for by the persistence of modes of presentation. Elsewhere, in the less prominent of his two

arguments, Evans (1985, p. 309) briefly makes this point a little more explicitly, and the following argument is at least partly inspired by his remarks.⁴

Let us recall the general structure of an argument for the persistence of a mode of presentation through time, as embodied in the Transparency Principle. If one thinks, at time t_1 , that 'Cicero was a great orator' then there is a kind of *ceteris paribus* rational obligation for one to predicate the same thing of Cicero, or at least not to predicate its negation, a moment later at time t_2 . There are, of course, plenty of circumstances in which this does not apply; one may forget one's earlier predication, for example, or one may simply change one's mind. But nevertheless the earlier predication *can* make the negation of the later predication irrational. This happens, quite simply, because at t_2 one remembers one's earlier thought that Cicero was a great orator. Having done so, and having accepted the earlier thought as true, one would be irrational to assert that Cicero was not a great orator. Through memory, one's thoughts at one time can thus have a rational bearing on what one should assent to at a later time. One's thinking can be diachronically transparent.

This diachronic rational relation implies the retention of a mode of presentation though time. Imagine that at t_2 one no longer possessed the term 'Cicero' but instead possessed a newly acquired term 'Tully' which referred to the same person referred to by 'Cicero' at t_1 but expressed a different mode of presentation. One would therefore find the identity 'Cicero = Tully' informative. Under these circumstances one's earlier thought about Cicero would have no bearing on what one should assent to about Tully at t_2 unless one also believed that 'Cicero = Tully'. In order to be able to entertain this identity, however, one would have to retain the mode of presentation 'Cicero' until

⁴ John Campbell (1987, 1994, pp. 73-88) gives a highly illuminating discussion of a similar argument for demonstratives, though he does not extend it to indexicals. Michael Luntley (1997, 1998, 1999) gives a variety of arguments for dynamic thoughts to the effect that the significance of a thought at one time can sometimes only be captured by acknowledging the rational bearing that it has on earlier thoughts, and Christoph Hoerl (1997) argues that the possibility of changing one's mind about whether something is predicable of an object requires there to be a continuous mode of presentation of the object throughout. The issue of cognitive dynamics is also related to a difficulty with demonstratives to which David Kaplan (1989a, pp. 588-590) draws attention.

one acquired the term 'Tully' at t_2 . A diachronic rational relation therefore implies the retention of a mode of presentation. The same argument can be applied to the modes of presentation expressed by the different occurrences of 'Cicero' at t_1 and t_2 . Consequently both occurrences of 'Cicero' must express the same mode of presentation. As an objection one could, of course, assert that the modes of presentation expressed by 'Cicero' at t_1 and t_2 differ just because of the time difference and that there is a brute rational relation between different modes of presentation at different times, but it is very hard to see a motivation for this view. Individuating modes of presentation in this way has no explanatory value.

An argument of exactly the same structure applies when there is a change of indexical terms. Consider a subject 'S' who acquires at time t_1 the belief that:

(3) It is rainy here

Suppose that S then moves away from the place at time t_2 but keeps track of it; this would typically be done perceptually, by keeping it in view. There are circumstances in which S's thinking (3) at t_1 would make it irrational for S to think about the same place at time t_2 that:

(4) It is not rainy there

The circumstances in question are those in which S remembers that the place is rainy and has no change of opinion on the matter.⁵ While keeping track of a place one has a constant epistemic engagement with it such that one never needs to reflect on whether it is the same place that one was thinking about earlier. Consequently even when it becomes appropriate to replace 'here' with 'there' there is the same kind of rational

relation between one's thoughts at the different times as exists between the earlier and later 'Cicero' thoughts described above. In both cases this can only be explained by a retained mode of presentation; the same mode of presentation is expressed by 'here' in (3) as is expressed by 'there' in (4).

It cannot be objected that the rational relation between (3) and (4) exists because of a suppressed identification 'here = there'. For such an identity statement would not be a possible object of thought. The mode of presentation expressed by 'here' is only available at time t_1 while the mode of presentation expressed by 'there' is only available at the later time t_2 ; as soon as the place becomes 'there' it ceases to be 'here'. To suggest that the modes of presentation would be available beyond these times would, of course, be to concede the point.

Neither can it be objected that the rational relation can be explained by an inference based on the memory that the place referred to earlier as 'here' is rainy, as follows:

- i. (From memory): the place that was referred to by 'here' at t_1 is rainy
- ii. There = the place that was referred to by 'here' at t_1 , therefore:
- iii. It is rainy there

Since the thought remembered at line (i) contains no demonstrative or indexical elements it can be assumed to be retained from t_1 to t_2 without begging the question. If the place thought about at t_2 was then judged, by whatever means, to be identical to the place that was thought about at t_1 (line (ii)) then it would be possible to infer the conclusion, line (iii).

This is certainly a way in which an initial thought 'it is rainy here' could lead to a later thought 'it is rainy there' and I do not deny that such reasoning often occurs. In such cases there is no reason to assume that 'here' and 'there' express the same mode

⁵ The adjective 'rainy' should be understood timelessly, so that if a place is rainy at one time then it is rainy at all times. If this seems counterintuitive it can be substituted for something unambiguously

of presentation. But an inference of this kind requires the place to be re-identified at t_2 . There may certainly be circumstances, those in which the thinker has failed to keep track of the place, in which such a re-identification is necessary. S might, for example, recognise the place by noticing certain distinguishing features which S remembers were possessed by the place seen earlier. But such reflective thinking differs from a common unreflective level of thought in which the thinker keeps track of the place, typically using an ongoing perceptual contact with it. This renders such re-identifications superfluous. We are concerned only with situations of the latter kind; it only these situations in which the mode of presentation is retained through a change of indexicals.

As a modified objection it might be suggested that it is by keeping track of the place that the thinker is able to make the re-identification; at t_2 S knows that 'there = the place that was referred to by 'here' at t_1 ' because the place has been kept track of throughout. But this would beg the question, for it implies that keeping track allows the thinker to know that 'there' satisfies 'x = the place referred to by "here" at t_1 ' simply by having known, at t_1 , that it was satisfied by 'here'. This assumes precisely the kind of diachronic rational relation that it was supposed to dispense with.

I have so far spoken of the subject keeping track of a place using perception. This typically involves using perception to keep track of objects which stand in stable relations to the place.⁶ There are, however, other ways of identifying places, though these only allow the subject to keep track of a place in relatively limited ways. It may, for example, be possible to identify a place 'egocentrically', making use of the spatial relation between the place and oneself.⁷ This way of identifying a place cannot be the norm because it would not allow one to keep track of places as accurately as we generally do. Imagine, for example, seeing a place in the distance and walking to it; it is not plausible that one could do this purely egocentrically, without the help of

timeless such as 'is SP's place of birth'.

⁶ Strictly speaking one could identify places using 'features' rather than objects; this refinement makes little difference, and I shall ignore it since the actual world contains objects.

perception. One could, however, keep track of a place egocentrically to the extent that, for example, if one is aware of moving a little to the left then one knows that the place must subsequently be identified in virtue of being a little to one's right.⁸

The distinction between different ways in which someone might keep track of a place is not important for our current purposes, however. For no matter how the place is kept track of, so long as it *is* kept track of there will be exactly the same kinds of rational relations between thoughts expressed using 'here' and 'there' at different times that have been discussed above. Consequently the argument does not rely on any assumptions about what is involved in keeping track.

The same argument can be applied whenever one indexical term can be replaced with a different co-referring term without the need for the thinker to re-identify the reference. Consider temporal indexicals. It is not possible to keep track of times perceptually in the same way that it is possible to keep track of places. A time can be kept track of as it recedes into the past, however, by keeping track of its location in a sequence of events which one can trace back from events taking place at the present moment. This is analogous to keeping track of places by touch, feeling one's way from one object to another and remembering their layout. It may also be possible to keep track of a time egocentrically over a limited period in a similar way to that described above for places. Keeping track of units of time such as days is more straightforward; one just has to keep track of the number of days which have passed. This would be difficult to do over long periods, but certainly the transition from 'today' to 'yesterday' presents no obvious problems.

⁷ The distinction between egocentric and demonstrative identifications of places is discussed in more detail in section 6.6, below.

⁸ This way of keeping track was also discussed earlier (section 3.2) as an example of a case in which we must either say that there is a continuous Way of Referring which has to be updated to keep up with the changing circumstances or else we must say that a referential episode does not rely on a strict overlap between different Ways of Referring. One can start to see here that the notion of 'keeping track over a period of time' cannot be sharply distinguished from the more general notion of a referential episode.

Indexicals which refer to persons do not lend themselves to dynamic transitions between different terms in the same way that spatial or temporal indexicals do. In particular, it is hard to imagine circumstances under which one could keep track of oneself while exchanging indexical terms between 'I' and, for instance, 'you'. There might, however, (with a little imagination) be ways of keeping track of someone who changes between 'he' and 'she'. In such cases the same arguments would apply.

5.3 Further considerations

The arguments just given can be backed up by a number of supplementary considerations. Consider, for example, what happens when a thought entertained at some earlier time is negated.⁹ Suppose I see a certain place in the distance and, thinking it looks attractive, I think:

(5) We should hold the barbecue there

I then walk over to the place, keeping track of it visually as I move towards it. On arriving at the place, however, I discover that it is not as attractive as it looked from a distance. I therefore change my mind and think:

(6) We should not hold the barbecue here

Now, the point of saying that I change my mind is that the later thought is not independent of the earlier one but rather has a certain rational bearing on it. Rather than simply thinking one thought followed by another, where the two thoughts are independent, my thinking (6) is inconsistent with my continuing to think (5). The best

way to explain this is to say that the 'there' of (5) and the 'here' of (6) express the same mode of presentation, and the change between (5) and (6) is just a change of predication.

The importance of modes of presentation which survive changes of context can also be seen in the role that they play in the justification of an on-going project over a period of time.¹⁰ During my search for a good spot for a barbecue I first see a location in the distance then go there to get a better look at it. My going and looking can be seen as a project. My decision to embark upon the project is taken in the light of the thought expressed by (5); I form an intention to go 'there' to get a better look. Now, suppose that when I arrived at the place the mode of presentation previously expressed by 'there' was lost and a new, independent one was expressed by 'here'. I would no longer have any grasp of what I was doing there. My being there would be part of an ongoing project (to go to the place and look around) which would require a continuous comprehension of my reasons for what I was doing. The mode of presentation expressed by 'there' would be likely to appear in a number of beliefs and desires which went together to form the background justification for the project (e.g. 'it looks nice *there* and I need to find a nice spot so I should go *there* to see if it is better *there* than the places I have seen so far' etc.) This project could only continue as I moved around if the mode of presentation survived the change of context.

Another point to consider is that since both 'here'- and 'there'-thoughts depend on the same ability to identify and keep track of places it seems arbitrary to claim that a mode of presentation of the place becomes unavailable when the linguistic expression changes. Imagine moving around relative to a particular place, sometimes passing through it then moving away again. Just the same ability to keep track is involved throughout. Imagine, for example, sitting on a train and seeing a place in the distance,

⁹ Cf. Luntley 1998; 1999, e.g. pp. 338-342.

¹⁰ Cf. Luntley, 1997.

passing through it and moving away again. Throughout the approach it would be possible to think:

(7) It is *G* there

It seems implausible that there would be a change in the mode of presentation of the place during the approach just because it gets steadily closer. After passing through the place a thought of the same form is possible and it seems similarly implausible to deny that the same mode of presentation of the place is retained as the place gets gradually further away. Now, at some point while passing through the place it becomes necessary to say 'here' instead of 'there'. This may happen very briefly. Given the continuity in the way in which the place is identified throughout the episode it would seem strange to suggest that the thoughts before and after passing through the place, both of which would be expressed using the word 'there', involve different modes of presentation. Indeed if the train had passed just a little to one side of the place instead of straight through it the thought would have been continuously expressed using 'there' and there would be no reason whatsoever to suppose that the mode of presentation differed depending on whether the train was moving towards the place or away from it. If all of this is correct then it seems arbitrary to deny that the same mode of presentation would also be available while the train was passing through the place, as well as both before and after. Consequently the same mode of presentation would be expressed first using 'there', then 'here', then 'there' again.

5.4 Diachronic anaphors

One final consideration depends on a certain understanding of the thoughts expressed by sentences containing anaphors. Imagine once again being on a train. As the train stands stationary at time t_1 I look around at the surrounding countryside and say:

(8) We should hold the barbecue here

The train pulls away, but I keep looking at the same place, never losing track of it. My continuing perception allows me to make further judgments about the place; from a distance I can see aspects of it that I could not see from closer up. At time t_2 based on these further perceptions I say:

(9) It is well sheltered from the wind

The word 'it' in (9) is an anaphor; it refers to the same place as the word 'here' in (8). Suppose that by the time I make the judgment expressed by (9) I am no longer at the place. Consequently it would no longer be appropriate for me to use the word 'here' in referring to it. Now, the most obvious account of anaphors says that the anaphor reproduces the mode of presentation expressed by the word it stands in for; in this case the word 'here' in (8). If anaphors work in that way then it follows that the mode of presentation expressed by the word 'here' in (8) is still available at t_2 , after I have moved away from the place.

Now, I could equally have said something equivalent at t_2 using the word 'there' instead of an anaphor. But we have just seen that the mode of presentation expressed by 'here' at t_1 is still available at t_2 . This makes it hard to deny that it is this same mode of presentation that would be expressed at t_2 by 'there'. The alternative would be to say that there would be two different modes of presentation of the place available

simultaneously: the one expressible using the anaphor and the one expressible using the word 'there'. This is highly implausible, especially since whatever is predicated of the place under one mode of presentation would have to be predicated of it under the other mode of presentation.

Although I do not have a simple argument to show that anaphoric thought has to be construed in the way just described it is a little hard to see what alternatives there could be. In the absence of any alternative account it is therefore difficult to deny that the same mode of presentation can be retained through a change of spatial indexicals. A similar argument could be given for other indexical terms insofar as it is possible to keep track of the reference through a change of indexicals, so the conclusion is quite general.

CHAPTER 6

The Role of Egocentricity in Indexical Thought

It has now been shown that the same mode of presentation can be retained through a change of indexical term. Indexicals, however, can very easily appear to be paradigm examples of terms which have characteristic inferential roles; roles by which the modes of presentation expressed could be at least partially individuated. If this were correct it would conflict with the conclusion just mentioned and would threaten Conceptual Atomism, which denies that concepts can be individuated in terms of inferential roles. The aim of this chapter is to show that this difficulty can be avoided.

6.1 *Indexical thought and action*

The difficulty just mentioned can be put in terms of a simple objection which has been surprisingly neglected by the advocates of dynamic thoughts. The objection is based on the fact that indexicals have special consequences for actions and inferences which differ for different indexicals. This has often been discussed in relation to communication, where it appears to present major problems for Frege's view that communication consists in the sharing of thoughts, and has also been discussed in other contexts. The problem for Fregean communication using indexicals is that when the shift of context between speaker and hearer requires that they use different indexicals it will often be rational for them to perform different actions. To borrow an example from John Perry (1977, p. 494), if you say that:

(1) A bear is about to attack me

you should roll up in a ball and try to keep still. But if I understand what you said I will express it using the sentence:

(2) A bear is about to attack you

This thought will give me reason to run and get help, not to roll up in a ball and try to keep still. The same phenomenon occurs with other indexicals. Consider the following two pairs of sentences, uttered in different contexts such that the indexical terms of each pair co-refer:

(3) Today is rainy

(4) Yesterday was rainy

(5) It is rainy here

(6) It is rainy there

The thinker of the first thought in each pair generally has a reason to put up or carry an umbrella, but the thinker of the second thought in the pair does not.

Now, if two people perform different intentional actions they must be in different overall psychological states. It does not, of course, follow immediately from this that the above pairs of sentences express different thoughts, for it could be that the thinkers perform different actions because they differ in thoughts additional to those expressed. But we can at least stipulate, for the purposes of argument, that there are no psychological differences that are unconnected with the thoughts expressed. For instance we could stipulate that you and I were in identical psychological states before the appearance of the bear. It follows that there is a difference in psychological state

associated with the difference in the indexical terms used; and it has seemed natural to many philosophers to suppose that this difference lies in the singular modes of presentation expressed by the indexical terms.

This view has led many philosophers to abandon the Fregean idea that indexical communication involves the sharing of thoughts. Some have suggested that while the speaker and hearer share a Russellian proposition they differ in the 'character' under which this proposition is entertained (Kaplan 1989). Character plays the role of mode of presentation in these theories but the Fregean one-to-one correspondence between singular modes of presentation and references is broken (Perry 1977). Others have stayed closer to Frege's original conception of thought, but have still rejected his account of communication. John McDowell, for example, suggests that much of Frege's account of thoughts could be rescued from these difficulties if indexical communication were seen as involving "different thoughts that, however, stand and are mutually known to stand in a suitable relation of correspondence" (McDowell 1984, p. 290). As mentioned above, Evans (1982, pp. 40, 315-316) had similar doubts about the Fregean account.¹ Many have also felt that the difficulties in the Fregean account of communication were at the root of Frege's rather problematic view of the first person, in which there is both a private and a public sense associated with 'I', the former being principally relevant to action while only the latter is communicable.

The distinctive psychological significance of an indexical term is not only manifested in communication. If, for example, yesterday, I thought 'today is rainy' it will have been rational for me, *ceteris paribus*, to take an umbrella with me when I went out. But if, today, I think 'yesterday was rainy' this gives me no reason to take an umbrella. Likewise, when I think 'it is rainy here' I may open my umbrella but when I move away from the place and think 'it is rainy there' I no longer have any reason to do this. Even the differences in the utterances of the different indexicals are deliberate and can therefore be regarded as different actions (different speech acts). Such differences in

rational actions require an explanation in terms of psychological states, and it has seemed natural to many philosophers to give this explanation in terms of differences in the modes of presentation expressed by the different indexical terms. Such an explanation only works if the different modes of presentation putatively associated with different indexicals have different inferential roles.

An example of this approach is found in the work of John Perry who notes the variation of rational action with indexical terms in the dynamic case as well as in communication:

As time passes, I go from the state corresponding to 'The meeting will begin' to the one corresponding to 'The meeting is beginning' and finally to 'The meeting has begun'. All along I believe of noon that it is when the meeting begins. But I believe it in different ways. And to these different ways of believing the same thing, different actions are appropriate: preparation, movement, apology (1979, p. 19).

When Perry speaks of 'believing the same thing' he does, of course, mean believing the same Russellian proposition, not the same Fregean thought. Perry's solution is to adopt a distinction similar to Kaplan's (1989) *kontent/character* distinction; in Perry's terminology, although 'what is believed' remains the same, the 'belief state' changes. In Kaplan's terminology it is the character which changes. This is, in effect, a change in a mode of presentation, though not one which is intrinsically related to the reference. As discussed above, the Perry/Kaplan theory is a version of Inferential Role Semantics.² But although ascribing different characters to the different indexicals offers a very neat explanation of the associated differences in rational actions it offers no resources to deal with the diachronic rational relations discussed in chapter 5. Like

¹ See above, section 4.1.

² See above, section 2.3.

other versions of Inferential Role Semantics, it buys variation in significance for rational actions and inferences at the cost of diachronic transparency.³

So there is a problem: Inferential Role Semantics can explain why thoughts expressed using different indexical terms differ in their characteristic significance for rational actions and inferences but cannot accommodate the retention of a mode of presentation through a change of indexicals as implied by the arguments of chapter 5. Conceptual Atomism, on the other hand, is compatible with the retention of modes of presentation through changes of indexical but owes us an account of why thoughts expressed using different indexical terms differ in their characteristic significance for rational actions and inferences. It is to this account that we must now turn.⁴

6.2 *The structure of indexical thought*

As a solution to the difficulty just mentioned I propose an account of indexical thought according to which there can indeed be a mode of presentation retained through a change of indexicals (and perhaps common to speakers in different contexts), but in addition there is a further mode of presentation which is associated with a particular indexical term and which therefore changes when the indexical term changes. It is this latter mode of presentation which accounts for the different psychological roles of different indexical terms. Let us call this the egocentric mode of presentation.

There are two main versions of this theory to consider. Firstly the egocentric mode of presentation may take the form of a subject term or secondly it may take the form of

³ Kaplan (1989, pp. 537-538) does note the difficulty, but leaves the matter unresolved.

⁴ Evans shows signs of being aware that there is a problem with the notion of a dynamic thoughts when he acknowledges that there are differences between different indexical terms which, it might appear, 'can be exploited to produce the possibility of differing epistemic attitudes to the thoughts, which would then preclude their being the same thought, if thoughts are intended to be the object of propositional attitudes' (1981, p. 292). His solution, however, is to point out that it is not possible to take differing attitudes to the thoughts in question *at the same time*. This does not seem an adequate

a predicate. It is the latter version which I think is correct. According to this theory, whenever there is a thought expressible using the word 'here' the thinker believes that the place in question has a certain property which I shall call the property of being here. Thus 'is here' is a predicate, of a kind which I shall call *egocentric* predicates (indicated by the underlining). The same applies, *mutatis mutandis*, for other indexicals. So, at t_1 , a thought expressible as 'it is rainy here' involves two different predications applied to the same singular term:

(7) p is rainy

(8) p is here

The letter ' p ' refers to the place and expresses a mode of presentation capable of being retained through a change of context and a change of indexical term. There are, of course, no such words as ' p ' or 'here' in English; they are merely being used as a means by which to indicate the structure of the thoughts in question.

Subsequently, at t_2 , when the thought becomes 'it is rainy there' there is a change of egocentric predicate but, if the place has been kept track of, everything else remains the same:

(9) p is rainy

(10) p is there

The retained mode of presentation expressed by ' p ' accounts for the transparency of thought over time and the change of egocentric predicate accounts for the differences in actions and/or inferences. A corresponding account of communication using

response; there are systematic differences between thoughts expressed using different indexicals which seem inconsistent with the subject being in the same overall mental state, albeit at different times.

indexicals, in which the speaker and hearer can share the mode of presentation ‘*p*’ but differ in the egocentric predicates applied, might also be possible.

Some general points must be made about this theory. Firstly, it need not be assumed that an utterance of ‘it is rainy here’ expresses two separate thoughts; instead of the separate thoughts (7) and (8), for example, there could be just one thought in which the two predicates (‘is rainy’ and ‘is here’) are attached to the same subject term, ‘*p*’.⁵ Secondly, from what has been said so far it is not necessary that there must be an egocentric predicate ‘is there’ at all; there could be an egocentric predicate ‘is here’ at t_1 and then no such predicate at t_2 . Much the same could be said about ‘now’ and ‘then’. Having registered this caveat, however, I shall continue to assume that there is an egocentric predicate ‘is there’, for when thinking of a place as ‘there’ one certainly has a sense of the egocentric location of the place. In fact there must be a potential infinity of different predicates corresponding to the infinity of different egocentric relations, and ‘is there’ should be understood as expressing whichever one is involved in the thought in question.

The alternative version of the theory, in which the egocentric mode of presentation corresponds to a subject term, might initially look equally promising. According to this version there are two different modes of presentation of the reference simultaneously, one of which changes according to the indexical term while the other can be retained. At t_1 , an utterance of ‘it is rainy here’ would imply thoughts of the following form:

(11) *p* is rainy

(12) **HERE** is rainy

Once again ‘*p*’ refers to the place and is capable of being retained through a change of context and indexical term. The word ‘**HERE**’ (when printed as shown) expresses

⁵ Such possibilities are discussed in greater detail below, section 6.5.

another mode of presentation of the same place, associated with the word 'here' but not with the word 'there'. As with the predicate theory there need not be two separate thoughts; instead of (11) and (12) there could, for instance, be just one thought '*p* is rainy and **HERE** is rainy' or '*p* and **HERE** are both rainy'. I shall not dwell on these possibilities, for all that is crucial to our current concerns is that there are two different modes of presentation of the place. At t_2 , given that the place has been kept track of, an utterance of 'it is rainy there' would imply a change of just one of these modes of presentation:

(13) *p* is rainy

(14) **THERE** is rainy

The retained mode of presentation '*p*' accounts for the diachronic transparency between t_1 and t_2 and the change of mode of presentation from '**HERE**' to '**THERE**' accounts for the variation in actions or inferences. It would also be possible to give an account of indexical communication according to which one singular mode of presentation, '*p*', is shared between the speaker and the hearer while the other is not; rather like Frege's account of 'I'.

To those who are used to certain standard theories of indexical thought the theory just described may, at a first glance, look tempting. For it may seem to offer the possibility of interpreting '*p*' as a perception-based 'demonstrative' mode of presentation leaving '**HERE**' as a 'purely indexical' mode of presentation, available even in sensory deprivation.⁶ As mentioned above, however, although we typically keep track of places using perception this is not a necessity. One can, to a more limited extent, keep track of a place without perceiving anything provided one can somehow keep track of one's own movements. If one knows that one has moved a little to the

⁶ Cf. Evans (1982, p. 161).

left, for example, then one knows that the place where one was must be a little to the right. The same diachronic rational relations exist as in the perceptual case and the same arguments apply. This means that there is no reason why '*p*' should always be related to perception; even in sensory deprivation there would have to be two modes of presentation of the place. So the theory does not incorporate a distinction between 'perceptual demonstrative' and 'purely indexical' modes of presentation and I suspect that this will make it appear a little less plausible.

In any case there is a more serious problem with this theory. The problem relates to how the thinker is supposed to get from thinking (11) and (12) to thinking (13) and (14). The transition from (11) to (13) is of course trivial; (13) is the same thought as (11), which is simply retained. The problem relates rather to (12) and (14). Why, for example, must someone who starts by thinking (11) and (12) at t_1 then go on to think not only (13) but also (14) at t_2 ? This is only explicable if there is an inference based on an identity of the form:

(15) *p* = **THERE**

By the same token a 'here'-thought must be associated with an identity of the form:

(16) *p* = **HERE**

If there were no such identities then there would be no reason to predicate the same thing (such as 'is rainy') of both '*p*' and '**HERE/THERE**'.

This would require the thinker of any 'here'-thought, for example, to judge that '*p*' is identical with '**HERE**'. But I can only see one way in which this could occur. Firstly, '**HERE**' would have to be identified by the thinker in virtue of uniquely having the egocentric property 'here'. Secondly there would have to be a judgment that:

(17) p is here

Thus, on the grounds that only one place can have the property 'here', it would be inferred that ' $p = \mathbf{HERE}$ '. In other words, the only way to judge that ' $p = \mathbf{HERE}$ ' would be to judge that the place ' p ' had the same egocentric relation to oneself that defines the place ' \mathbf{HERE} '.

The trouble with this idea is that it renders ' \mathbf{HERE} ' and ' \mathbf{THERE} ' redundant. For (17) is exactly the same thought as (8); it involves an egocentric predication exactly as the predication theory says. But if the egocentric predication is necessary anyway then there is no longer any need to posit the modes of presentation ' \mathbf{HERE} ' or ' \mathbf{THERE} '. They have no explanatory value. Hence the analysis in terms of egocentric predicates is the correct theory.

6.3 *Egocentric Predicates*

Egocentric predication has so far been described as a technical notion introduced in order to solve a technical problem but in fact some egocentric predicates are intuitively quite familiar. The egocentric predicate components 'here' and 'there', for example, can be thought of as similar to the adjectives 'near' and 'distant', which are also egocentric predicates. The egocentric predicates associated with indexical thoughts belong to a more general category of egocentric predicates, defined as those predicates which pick out a relation in which the thinker is one of the relata.⁷

When thinking about a place one can be aware of where the place is in relation to oneself. This is a property of the place; a relational property, but a property

⁷ The following account of egocentric predicates has benefited from John Campbell's (1993a, pp. 82-88, 1994, pp. 41-46) account of *causal indexicals*, with which egocentric predicates have much in common.

nonetheless. It is this property which I suggest is predicated of a place by a spatial egocentric predicate. There are, however, two quite different ways in which this property could figure in thought. It may, firstly, be conceived of by the thinker as relational. The place is thought of as being 'where I am', 'near to me', 'distant from me', and so on. There is no doubt that such relational thoughts occur, and are often relevant to decisions about which actions to perform. They are also sometimes essential for communication with others.⁸

It is not obvious, however, that spatial thought must always be conceived of by the thinker in this explicitly relational way. Moreover, the equivalent relational predicates in the temporal or personal cases arguably fail to capture the nature of the thoughts normally expressed by temporal or personal indexical terms. In the temporal case since the thinker exists at many different times the relational egocentric predicates would have to contain a reference to something other than the thinker, perhaps to the time at which the thought itself is occurring. But although such thinking certainly occurs (and is required, for example, in understanding other speakers' sentences uttered at times other than the present) it is not clear that unreflective indexical thought about times is normally understood by the thinker as relational.⁹ It is even less plausible that the corresponding predicates for personal indexicals are always relational. In the first-person case the relational predicate would be of the form 'is the thinker of this thought'; it is not plausible that first-person thought depends on such articulations.

When thinking of a place as near or distant it is not necessary to think of this property explicitly in relation to oneself. Perhaps, then, at the most primitive level of

⁸ There is, however, some question as to whether it would be intelligible to claim that what it is for a place to be 'here' could ultimately be understood in terms of its relation to one's own position. See Evans 1982, pp. 153-157 and Campbell 1993a, pp. 71-76, 1994, pp. 8-16.

⁹ Higginbotham's (1995) account, in which part of the content of certain thoughts is the thought itself, has some broad similarities with the relational account of egocentric predicates. Higginbotham cites examples of sequence of tense and other linguistic phenomena which illustrate the way in which the logical form that accounts for sequence of tense (on his account, a temporal relation between the utterance and a 'situation') must sometimes be part of the thinker's mental state. I need not disagree with this; such reflective uses of indexical terms must be possible. The question is whether they are universal.

thought the egocentric predicates associated with the use of indexicals could be monadic; although they refer to a relation involving the thinker, they are not conceived of by the thinker as relational. Instead, they are unstructured predicates whose use is very similar to the unstructured predications 'is near' and 'is distant'.

In the temporal case, the idea that predicates like 'is past' can be understood monadically seems particularly plausible, especially given that many people (including a number of philosophers) find it natural to assume that properties like pastness or futurity, far from being relational, are in fact intrinsic properties of times.¹⁰ If the theory is also applicable to the first person then the egocentric predicate associated with 'I'-thoughts would be a monadic predicate 'is me', which predicates identity with a certain person (the thinker), but is not conceived of in this way by the thinker.

6.4 The role of egocentric predicates in communication

As noted above, many philosophers have abandoned the notion of indexical communication as involving the sharing of singular modes of presentation. On the theory that has now been outlined, however, there is no longer any need for this conclusion because the differences in actions between speaker and hearer can be accounted for in terms of differences in egocentric predication rather than differences in the singular mode of presentation. The theory of communication developed in chapter 4 can therefore be applied to indexicals; so long as both speaker and hearer understand the general rules of use for indexical terms (e.g. that 'here' refers to the location of the speaker) then it is possible for there to be the kind of mutual deference

¹⁰ If the advocates of this view (the 'tensed'- or A-theory of time) were correct then temporal predicates like 'is past' should be regarded as non-egocentric monadic predicates. There is still an active debate on this issue, much of which revolves around the nature of temporal indexicals. For an argument against intrinsic tenses which is independent of such issues, however, see Appendix C.

required for them to share in a single referential practice. But what role do egocentric predicates play in communication, and can they themselves be communicated?

Anyone who can understand an utterance containing an indexical term must understand the way in which the term combines with context to determine the reference, and must also assume that the speaker understands the functioning of the term in the same way. This provides us with a clue concerning the role of egocentric predicates in communication. Take the word 'here', for example. If you understand my use of 'here' then you must understand not only that I am referring to my own location but also that since I am expressing my thought using this term I must believe, of that particular place, that it is where I am. I must, in other words, have a belief of the form '*p* is where I am'. This involves the application of a relational egocentric predicate, just as the theory suggests. Similarly, if I utter a sentence containing the word 'past' then I must believe, of a certain time, that it is prior to my utterance and if I utter a sentence containing 'I' then I must believe that it was I who uttered it.

Each indexical utterance therefore implies that the speaker predicates, of the reference, the relation picked out by the relevant relational egocentric predicate. Moreover, since the same words can be uttered by different speakers and in different contexts, both speaker and hearer must understand that each relational egocentric predicate is an instance of more general relation such as '*π* is where *s* is' (where '*π*' and '*s*' are variables standing for a place and a thinker respectively). Without this understanding communication would be impossible.

This does not, however, preclude the possibility that an understanding of relational egocentric predicates could be dependent upon or additional to an understanding of the corresponding monadic predicates. It is important, in this connection, to bear in mind the difference between entertaining a thought which could be expressed using certain words and actually so expressing it. It is only in the latter case that it is necessary to insist that the thought involves the relational egocentric predicate. It may be, then, that in understanding someone else's indexical utterance one can, via an understanding of

the relevant relational egocentric predicate, put oneself in the other person's position by understanding what kind of monadic predication is involved in their thinking. It is in this way that one can literally come to understand another person's point of view.¹¹

If it is assumed, with Frege and most 'neo-Fregean' philosophers, that a mode of presentation determines a unique reference then monadic egocentric predicates cannot be shared by different people. Only one person can ever be in a position to make judgments involving a given monadic predicate because the predicate picks out its reference in virtue of the fact that it is entertained by that particular person. In judging that a place is 'here', for example, I have only to focus my attention upon the place; I do not have to pay attention to my own location as well. Only I can make the judgment in this monadic way. The nearest anyone else could come to this judgment would be to pay attention to both the place and to me, and to make a judgment about the relation between one and the other. Consequently only I can entertain the particular monadic predicate associated with my 'here'-judgments. This is consistent with the account of communication that has been given, however, and there is nothing obviously absurd about the idea that there are thought components that one can entertain only by being a particular individual.

6.5 Indexicals and complex demonstratives

There is an analogy between the dynamics of indexicals and the dynamics of complex demonstratives which, I hope, will help to clarify the role of egocentric predicates within the theory that has been outlined. Complex demonstratives are those in which the demonstrative term is modified by a nominal, as in 'that bottle'. To see the analogy

¹¹ These remarks have been influenced by Campbell 1998, pp. 130-131. Campbell argues that there must be a 'two-stage construction' involved in understanding others, involving both monadic and relational notions in order to picture what it would be for one's own monadic notions to be someone else's.

with indexicals, imagine sitting in front of a bar behind which there is a row of bottles on a shelf. One of the bottles, the tallest one, happens to be the only empty one. It would be possible to think 'that empty bottle is the tallest'. Now suppose that the bottle is slowly filled. It would be possible to keep track of the bottle perceptually throughout. But the expression of the thought would now change through a modification of the nominal: 'that half-empty bottle is the tallest' then 'that almost-full bottle is the tallest' and so on. These differences in expression could also lead to differences in actions or inferences if it happened to be important to the thinker how full the bottle was. Yet if a mode of presentation of an object can be retained by keeping track of the object perceptually it should be possible to retain the mode of presentation of the bottle while it is filled.

It is not difficult to make sense of this example. Although the same mode of presentation of the bottle is retained throughout there is a change in what is predicated of it as it is filled. The question arises whether 'that empty bottle is the tallest' and 'that half-empty bottle is the tallest' express the same thought. The answer depends on the view one takes of complex demonstratives. Many philosophers have held the view that the nominal does not contribute to the truth conditions of a sentence containing a complex demonstrative, and this could imply that sentences of the form 'that is *G*', 'that *F* is *G*' and 'that *H* is *G*' could all express the same thought. Opinions vary among these philosophers, however, as to what the nominal does.¹² Some hold that it has a restrictive role such that 'that *F*' can only refer to an object which is *F*. Others think it has only a pragmatic role, helping to direct the hearer's attention to the object. On this pragmatic interpretation an utterance of the form 'that *F* is *G*' indicates the presence of two thoughts: 'that is *G*', which is the thought expressed and retained through changes in the nominal, and 'that is *F*', which is not expressed by the sentence, though the fact that it is among the speaker's beliefs is manifested by the choice of words. The latter thought is exchanged for a different one when the nominal changes.

I hope the intended analogy is clear: as with indexicals, complex demonstratives involve a combination of a singular mode of presentation which can be retained through certain changes with a predicate which can change. In the case of complex demonstratives what changes is the nominal; in the case of indexicals, I have suggested, it is an egocentric predicate.

Ernest Lepore and Kirk Ludwig (2000) have recently claimed that the nominal does in fact contribute to the truth conditions of sentences containing complex demonstratives. They argue that the concatenation of a demonstrative with a nominal is a form of restricted quantification so that 'that *F* is *G*' shares interpretive truth conditions with a sentence of the form "the *x* such that *x* = *that* and *x* is *F* is such that *x* is *G*".¹³ This contains a pure demonstrative, '*that*', but the nominal, '*F*', is indirectly predicated of it via the quantification and therefore contributes to the truth conditions.

If this account is correct it may have implications for indexical thoughts as well. For it suggests a possible analogous theory according to which instead of pairs of thoughts like (7) and (8) there is a single thought associated with 'it is rainy here' which could be paraphrased along the lines of: "the *x* such that *x* = *p* and *x* is here is such that *x* is rainy".¹⁴ Whether or not this is correct does not affect the fundamental claim of the theory of indexical thought for which I have been arguing, which is that the difference between thoughts expressed using different indexical terms is a difference in

¹² See Lepore and Ludwig 2000, pp. 200-201, for a useful survey of the literature.

¹³ Lepore and Ludwig also give a more precise formulation in terms of satisfaction, suited to a Davidsonian semantics. I have paraphrased their formulation of the sentence above into English to avoid the need to explain their terminology. Lepore and Ludwig's (p. 200) primary concern is the semantics of complex demonstratives, and they do not make explicit the implications of their discussion for the thoughts expressed using these terms.

¹⁴ For the purposes of constructing a semantic theory something more complex may be required because '*p*' is not actually a lexical item (cf. Lepore and Ludwig, p. 237-238). My purpose here, however, is only to consider what kinds of thoughts might be expressed. In the appendix to their article Lepore and Ludwig (pp. 236-238) briefly raise the possibility of extending their account to more specialised terms. Using 'he' as their example they raise the possibility of an analysis in which the predicate 'is male' is involved in a way analogous to 'is *F*' in the case of complex demonstratives. They do not discuss the 'here' case explicitly but presumably it would involve a predicate 'is here'. For Lepore and Ludwig, whether such an account should be adopted turns on what one can infer from sentences containing the term in question (whether, for instance, a true or false assertion of 'he is

predication and not necessarily a difference in the singular mode of presentation expressed. It is, however, a version of the theory worth noting.

6.6 *Immunity to error through misidentification*

Certain judgments expressible using demonstrative or indexical terms are commonly said to be immune to error through misidentification. There are, I shall suggest, two different kinds of immunity to error through misidentification (henceforth 'IEM'), either of which can be exhibited by an indexical judgment.¹⁵ This is most naturally explained by the theory of indexical thought for which I have been arguing and thus provides an additional argument for it.

IEM is usually defined along the following lines: a judgment of the form '*a* is *F*', where '*a*' stands for an object and '*F*' stands for a property, is immune to error through misidentification if it makes no sense, having judged that *something* is *F*, to wonder whether it is *a* that is being judged to be *F*. When I have an empty feeling in my stomach, for example, it makes no sense for me to judge, on that basis, that someone is hungry and yet to wonder whether the person in question is me. My judgment that I am hungry is therefore immune to error through misidentification. This immunity is a property of judgments, not of sentences or the thoughts expressed by sentences. I may, for example, judge the same sentence, 'I am hungry', to be true as a result of being shown an X-ray photograph of an empty stomach. But this judgment could be in error through misidentification: the person whose stomach appears in the X-ray might not be me.

handsome' implies that something is male). The arguments given here provide an independent motivation for adopting such an account.

¹⁵ This is not, however, the same distinction made by Andrea Christofidou (2000).

Much of the interest in IEM has been due to the fact that many first-person judgments exhibit the phenomenon.¹⁶ This suggests that thinking about oneself does not always require identifying oneself in any fallible way and this raises interesting questions about the nature of first-person judgments. IEM does, however, occur in judgments involving other kinds of terms. To illustrate the two different kinds of IEM it will help to consider complex demonstratives once again. I shall argue that complex demonstratives exhibit a different kind of IEM from judgments expressed using terms of the form ‘the *F*’, ‘that (the *F*)’, or any other term whose reference must be identified by its properties. I shall then discuss the relevance of this distinction to indexicals.

Suppose, then, that I see an angry-looking wasp. I may judge that:

(18) That wasp is angry

My ability to identify the wasp rests on the fact that I receive a stream of information from it through perception. I have no grip on the object of my judgment other than via this information link. So, since my judgment about the angeriness of the wasp also rests on the same information, there is no possibility that the object which I judge to be angry is not identical to the one which I refer to as ‘that wasp’. My judgment is therefore immune to error through misidentification because I cannot judge, in this way, that something is angry and yet wonder whether it is that wasp which is angry. Borrowing an expression from Evans (1982, p. 180), we can describe judgments which exhibit IEM for this reason as *identification-free*.

This contrasts with a situation in which, upon seeing an angry-looking person, I judge that:

¹⁶ Recent discussion of the phenomenon originates from Wittgenstein (1958, pp. 66-7) and Sydney Shoemaker (1968, 1970). Evans (1982, sections 6.6 and 7.2) also gives an important discussion.

(19) Smith is angry

This judgment is only correct if the person I see is in fact Smith. It therefore rests on the truth of an identity between ‘that person’ and ‘Smith’, where the former refers to the person seen. If this identity is false then the judgment is in error through misidentification. Let us therefore call such judgments *identification-dependent*.¹⁷

Some judgments which are identification-free can, however, be in error through a different kind of misidentification. Suppose, for example, that the insect which I saw was not a wasp, but a bee. My taking it to be a wasp could reasonably be described as a form of misidentification. The error in this case, however, is not in an identity but in a predication; the insect is incorrectly judged to be a wasp.

The kind of misidentification just described can be made more vivid by replacing the structured complex demonstratives ‘that wasp’ and ‘that bee’ with unstructured terms ‘thasp’ and ‘thee’ respectively. These simply abbreviate the complex demonstratives. No such terms exist in English, of course, but there is nothing to prevent them from being introduced. On mistaking the bee for a wasp I would therefore make the judgment:

(20) Thasp is angry

Had I correctly identified the insect as a bee I would have judged instead that:

(21) Thee is angry

The error in (20) is a misidentification of ‘thee’ for ‘thasp’. The fact that (20) is identification-free does not prevent this kind of misidentification from occurring.

¹⁷ The identity in question need not always be articulated; it might be possible to judge directly that Smith is angry, but the truth of one’s judgment would still rest on the truth of an identity.

Some judgments are immune to the kind of misidentification just described. Suppose that I acquire the true belief that there is just one wasp and one bee in the room with me. Upon seeing an angry-looking insect I might then make a judgment that:

(22) The wasp is angry

Given the background knowledge that there is only one wasp in the room a unique reference can be picked out by the expression 'the wasp'. Unlike the reference of 'that wasp' the reference of 'the wasp' cannot fail to be a wasp because it is picked out by virtue of being the one and only wasp in the room. The second kind of misidentification therefore cannot occur. This judgment, however, is identification-dependent. For suppose that once again the insect which I saw was the bee rather than the wasp. The expression 'the wasp' would not refer to the insect which I saw and my judgment would therefore be in error through misidentification.

The terms 'thasp' and 'thee' could be given a dual use; as well as abbreviating complex demonstratives they could also be used to abbreviate 'the wasp' and 'the bee' respectively under the circumstances in which there is only one wasp and one bee present. When they are used in this latter way the judgment:

(23) thasp is angry

has the same properties as (22); it is identification-dependent but is immune to the second kind of error through misidentification because the associated predicate 'is a wasp' has a stipulative rather than informative use. This is the exact opposite of the

situation in which the word 'thasp' abbreviates the complex demonstrative 'that wasp'.¹⁸

Now, as implied above, it is not only judgments containing complex demonstratives or descriptions which can possess the two different forms of IEM. Any judgment whose object is identified using the same information upon which the judgment is based is identification-free and any judgment whose object is identified by virtue of a property which it possesses may be immune to the second form of error through misidentification if the referring term has an intrinsic association with the property.

Let us therefore consider judgments expressed using the word 'here'. I suggest that there are two different ways in which the place referred to by 'here' can be identified by the thinker. I shall call these *demonstrative* and *egocentric* identifications. They are analogous to the ways of identifying the reference associated with 'that wasp' and 'the wasp' respectively and correspondingly 'here'-judgments can exhibit two different kinds of IEM.

Let us consider demonstrative identifications first. This means that the Way of Referring to the place is demonstrative.¹⁹ Suppose, for example, that I see a wasp in my vicinity. Assuming that the word 'here' can be used to refer to a region of space which extends beyond the boundaries of my body, I might judge that:

(24) There is a wasp here

¹⁸ There are some connections between the distinction made here and the referential/attributive distinction made by Donnellan (1966) with regard to definite descriptions. Donnellan's two versions of 'Smith's murderer is insane' correspond to one case in which the reference is identified in virtue of fitting the description and another case in which it is identified perceptually and believed (mistakenly) to be the item which fits the description. In the former case fitting the description is non-negotiable but a judgment of identity with a perceived object would be informative. In the latter case the identity with the perceived object is non-negotiable but fitting the description is not.

¹⁹ A demonstrative identification should not be confused with a use of the word 'here' as a demonstrative as in the sentence 'Smith is *here*', said while pointing at a place on a map. On the contrary, all of the uses of the word 'here' which I shall discuss are bound by the rule that the word 'here' refers to the place where the speaker is (or the thinker, if there is no utterance).

When the identification is demonstrative the place referred to by 'here' is identified perceptually using information received from any objects within the region referred to as 'here', including the wasp itself. Judgments involving demonstrative identifications of places are identification-free for exactly the same reason as judgments about 'that wasp'. The stream of perceptual information upon which the judgment is based is also used to identify the place which the judgment concerns. Consequently, when 'here' is identified demonstratively it makes no sense to judge, based on the information received from the place, that there is a wasp somewhere but to wonder whether there is a wasp 'here'.

As with judgments expressed using 'that wasp' or 'thasp', however, identification-freedom does not render 'here'-judgments involving demonstrative identifications immune to a second kind of error through misidentification. In referring to a particular place there is a fact of the matter about whether it is appropriate to use the word 'here' rather than, for example, 'there'. Using the wrong term would be an error, one which could reasonably be described as an error of misidentification. This contrasts with simple demonstratives like 'that' whose use is much more flexible.²⁰

Now, mistaking 'thasp' for 'thee' (in their demonstrative use) would involve mistaking the property of being a wasp for the property of being a bee. Mistaking 'here' for 'there', I suggest, would also involve an error in ascribing a property, but this property is relational. The property is, of course, the one expressed by the egocentric predicate associated with the indexical term. One uses the word 'here' in referring to a place when one judges the place to be here, just as one would use the word 'thasp' in referring to an insect when one judged it to be a wasp. An error through misidentification of the second kind would result if a demonstratively

²⁰ There may be a hint at the corresponding point for the first-person in Sydney Shoemaker's (1968, pp. 558-9) pioneering discussion. Having noted that demonstratives exhibit the phenomenon of identification-freedom to which Evans (1982) later drew attention, Shoemaker suggested that this is not the same phenomenon that applies to the first-person. He based this claim upon certain differences between demonstratives and the first-person, the main one of which is that whereas a demonstrative does not have its reference fixed by any rule the reference of the word 'I' is fixed for a given speaker.

identified place was judged incorrectly to be here; the word 'here' would then be used in error in referring to it. Although rare in practice, this kind of error can indeed occur. If, in the example above, there was a distortion in the air between me and the wasp it would be possible for the wasp to appear to be much nearer to me than it actually was, due to the same kind of refractive effect that makes a swimming pool look shallower than it really is. I would identify the place demonstratively but I would get the egocentric predication wrong. Consequently if I judged that 'there is a wasp here', although I could be correct in judging that there was a wasp at the place which I identified demonstratively using information received from objects including the wasp I could be incorrect in referring to that place as 'here'.

Egocentric identifications differ from demonstrative identifications in the fact that they involve stipulating, rather than judging, that the egocentric predicate applies. In other words, the place referred to as 'here' is identified using either a descriptive or recognitional Way of Referring by virtue of being the place which uniquely possesses the egocentric property of being 'here'. This corresponds to the way in which the property of being a wasp is used to identify the reference of 'the wasp' or 'thasp' in the second kind of example discussed above. We need not, of course, assume that the thought is descriptive in form just because the Way of Referring may be descriptive (similarly there is no need to assume that 'thasp'-judgments must be descriptive; 'thasp' could be a directly referring term involving a descriptive Way of Referring).

I suggest that just as the word 'thasp' can be used in expressing judgments involving either of the two different ways of identifying its reference, so the word 'here' can be used to express judgments involving either demonstrative or egocentric identifications of a place, depending on the situation. Demonstrative identifications, for example, allow places to be kept track of while the thinker moves about. They allow someone to look at a place some distance away, wonder whether it is warm there, then go to the place using an ongoing perceptual information link to home in on it, and finally think 'ah yes, it is warm here'. An egocentric identification, on the other hand,

allows someone to entertain 'here'-thoughts whilst in sensory deprivation. This duality does not, however, imply a dual linguistic function of the word 'here'. No matter how the thinker identifies the reference, the use of the word 'here' is still governed by the rule that it refers to the place where the speaker is.

In case it should appear tempting to think that the identification of 'here' is always egocentric and that what I have been calling a demonstrative identification of the place referred to as 'here' is really a case in which a demonstratively identified place is judged to be identical to the egocentrically identified place referred to as 'here', recall the response to an equivalent objection given above.²¹ In order to judge that the demonstratively identified place was identical to the egocentrically identified place it would be necessary to judge that the demonstratively identified place was here; the two places would then be identified because they both possessed the same egocentric property. But in that case the demonstrative identification would have to go together with an egocentric predication anyway and this makes the assumption that there must always be an egocentric identification superfluous. In any case, if the objection were correct then 'here'-judgments would never be identification-free (given that, as I shall argue, judgments based on egocentric identifications are identification-dependent), yet there is surely a strong intuition that in a great many cases the word 'here' refers to a region of space which is identified demonstratively without any need for an identification with a place identified in some other way.

Now, when a place is identified egocentrically judgments about it are immune to the mis-ascription of egocentric predicates because the place is identified by virtue of the fact that it has the egocentric property. Such judgments are thus analogous to judgments about '*the wasp*' rather than '*that wasp*'. By the same token, however, they are not identification-free. The perception of a place normally affords an independent demonstrative identification of it even when the reference of 'here' is determined

²¹ See section 6.2, above, in which the 'subject-term' version of the theory of indexical thought was refuted.

egocentrically. Consequently, if there is a refractive illusion like the one described above the egocentrically identified place will not be identical to the demonstratively identified place, even though both places appear to the thinker to have the egocentric property 'here' (one place will have this property, the other one will not). Error through misidentification can therefore occur. There are, of course, certain cases in which the perception of a place does not afford a genuine demonstrative identification of it at all; for example when the subject is unknowingly looking at an image of a place relayed through virtual reality spectacles. In such cases it is not possible for the judgment to be in error through the falsity of an identity statement 'that place = here'. Even in these cases, however, it makes sense to judge that there is a wasp *somewhere* but to wonder whether there is a wasp *here*.

Evans (1982, pp. 179-191), in his discussion of IEM, places a restriction on the phenomenon such that it only occurs when the subject is taking it for granted that the perceptual information upon which the judgment is based is being received in the 'normal' way, not through unnatural media such as virtual reality spectacles or, in Evans's example, undetected headphones which relay sounds from another place. In this section of his work Evans seems to assume that 'here'-judgments always involve what is, in effect, an egocentric identification of the reference yet he nevertheless claims that such judgments are identification-free. This, if I understand correctly, is because he takes it for granted that an egocentric identification of a place always correctly picks out the place from which the perceptual information is received except in cases in which the information is not received in the 'normal' way. But as we have seen, even in 'normal' cases there can be two independent identifications of a place, one egocentric, the other demonstrative, which can come apart when there are refractions.²² It follows that in normal cases judgments involving an egocentric

²² Evans's error is hidden by the fact that his examples involve proximal sensory modalities such as touch, for which no perceptual distortion analogous to refraction is possible.

identification of a place are never identification-free. There is always some identity of the form:

(25) here (identified egocentrically) = here (identified demonstratively)

such that the identity could in principle be false, and if it were false the judgment about the place would be in error through misidentification. Evans tries to exclude cases in which the identity would be false by definition; but if this definition were extended to cover even cases of refraction the resulting claim about identification-freedom would surely be a rather trivial one (error would simply be being ruled out by definition).

To summarise, judgments expressible using the word 'here' can be based on two different ways of picking out a place, and to these there correspond two different kinds of IEM. 'Here'-judgments involving a demonstrative identification have in common with 'that wasp'-judgments that they are identification-free but not immune to a kind of misidentification which involves the mis-ascription of a certain predicate. 'Here'-judgments involving an egocentric identification, by contrast, are analogous to 'the wasp'-judgments in that they are immune to errors in the predication, but if they are based on perceptual information they are not identification-free.

Let us now consider whether the conclusions drawn for 'here'-judgments can be extended to other indexical terms. With regard to temporal indexicals it is possible, at least in principle, to distinguish demonstrative identifications of a time from egocentric identifications, although it has to be admitted that demonstrative identifications probably occur very rarely in practice. Whereas demonstrative identifications of places depend on the perception of objects demonstrative identifications of times depend on the perception of events. Most perceived events occur at the present time. Consequently demonstrative identification can normally occur only with indexicals which refer to the present time, such as 'now'. A time which is identified

demonstratively via a witnessed event is almost always correctly judged to have the property expressed by the egocentric predicate 'is now'.

This is a consequence of the fact that one's perceptions normally concern events in one's vicinity. When an observed event is further away, however, the time that the information takes to travel becomes a factor. Imagine looking into the night sky and seeing a star explode. The explosion may have taken place many years previously but someone who was unaware of the time that light takes to travel across interstellar distances might think that the explosion was occurring at the present time. This is very much like the case of refraction discussed above for spatial indexicals. Someone in this position might judge that:

(26) A star is exploding now

If, however, the time was identified demonstratively via a demonstrative identification of the event then the associated egocentric predication 'is now' would be in error. If, on the other hand, the time was identified egocentrically (by virtue of being uniquely now) then the judgment would be in error through a misidentification of the time (such judgments, as in the spatial case, are not identification-free but are susceptible to errors in which the egocentrically identified time is not identical to the demonstratively identifiable time at which the event occurs).

Let us now consider the first-person. I cannot possibly do justice here to all of the issues that the first-person raises and I shall have to leave many questions unanswered. It is, however, possible to make some tentative suggestions based on what has been said about 'here'- and 'now'-judgments. Let us assume that there are judgments in which the first-person cannot be used in error. Now, no one else can use the first-person correctly in referring to me; therefore no one else can have access to a way of making a judgment about me such that the first-person cannot be used in error. We

must therefore consider whether there can be a way of making a judgment about a person which is only available to that person.

There is relatively little information about a person which is available only to that person. There are ways in which other people can find out about one's physical states and in many cases one's mental states as well. There are, however, *ways* of gaining information about oneself that are not available to other people. If I judge that 'I have an empty stomach' there are various ways in which someone else could also acquire the information that I have an empty stomach. But whereas anyone could acquire this information by looking at an X-ray only I could acquire it by feeling that my stomach was empty. It is, of course, apparent to me whether or not I am acquiring information about myself in one of these privileged ways; I perceive that someone's stomach is empty but there is something about the way in which the information is received that differentiates the person in question from all others. It therefore seems reasonable to suppose that receiving information in a privileged way leads me to predicate something of the person about whom the judgment is made. This predication is, of course, egocentric; there is an egocentric predicate 'is me', which picks out the relation of identity between the person referred to and myself, though I would have to understand the predicate as monadic in order to avoid circularity.

Those judgments which are made on the basis of a privileged way of acquiring information are, as a matter of fact, the ones which are commonly held to be immune to error through misidentification relative to the first-person. Now, if it were impossible ever to receive information about another person through one of the privileged channels just mentioned then it would indeed be impossible for a sincere use of the first-person to be inappropriate. It should be noted, however, that this is consistent with either a demonstrative or an egocentric identification of the first-person. Each type of identification would make one type of error through misidentification logically impossible but would leave the other kind of error possible at least in principle. Consider demonstrative identification. It is perhaps debatable

whether information received in the privileged ways mentioned above always affords a genuine demonstrative identification of a person. If it did afford such an identification, however, then the judgment would be identification-free. The thinker would judge the person identified to have the egocentric property of being 'me' as a result of the way in which the information was acquired. But if the information actually came from someone else then the predication would be in error.

On the other hand, and perhaps more plausibly in most cases, the person could be identified egocentrically by virtue of satisfying the egocentric predicate 'is me'. This would make it impossible to mis-ascribe the egocentric predicate but the judgment would not be identification-free. It would be susceptible in principle to error through the falsity of an identify between the egocentrically identified thinker and the demonstratively identifiable source of the perceptual information who appeared, falsely, to possess the egocentric property 'me'. Since in practice information received in the privileged ways normally concerns the thinker, error through misidentification does not normally occur. But this is not a logical immunity to error; it is an immunity which rests upon a contingency about the way in which information about people is acquired.

6.7 I am here now

Among the reasons why Kaplan drew the content/character distinction was the need to be able to deal with a phenomenon illustrated by the sentence:

(27) I am here now

Kaplan (1989, p. 509) claims that anyone who understands this sentence knows that it cannot be uttered falsely; it can, in other words, be known a priori. Yet the following sentence (where the symbol '□' implies necessity (truth in all possible worlds)) is false:

(28) □ I am here now

It is a contingency that I am here now; I could have been somewhere else. It must therefore be explained how it is possible to know a contingent truth a priori. The kontent/character distinction offers some apparatus with which to deal with this. It is the character of the terms which make up (27) that makes it knowable a priori, for the combination of characters is such as to always pick out a true kontent (the combined character can, of course, pick out different kontents in different contexts but they are always true kontents). Yet the kontent thus picked out is the same as that of certain other sentences such as:

(29) SP is in Coventry at 2.00 pm on 17th July 2001

This sentence is clearly contingent. So on Kaplan's account it is the kontent which is contingent and the character which makes the truth of the sentence knowable a priori.

I have rejected Kaplan's kontent/character distinction insofar as it applies to thoughts. I need not reject it as a purely linguistic matter. It would be possible, for example, to question why the sentence (27) always expresses a truth given that the truth in question is contingent. Kaplan's approach may be a perfectly acceptable way of answering this. But if (27) is knowable a priori this must be accounted for at the level of thought. Consider, for example, someone who utters (27) without understanding its meaning. Such a person has uttered a truth just in virtue of the linguistic characters of the terms from which the sentence is composed. But such a

person does not understand the thought expressed by (27) and therefore cannot know this truth a priori. The latter must be explained at a different level.

Now, since I reject Inferential Role Semantics I cannot account for the a priori nature of (27) in terms of the inferential relations between the different indexicals. I certainly cannot account for the a priori nature of (27) in terms of it being analytic because I reject analyticity. I can, however, account for it by considering the role of egocentric predication in (27). There are two cases which we must consider: demonstrative and egocentric identifications of the reference.

Firstly, suppose that the references of 'I', 'here' and 'now' were identified demonstratively.²³ I have already admitted that this is likely to be only rarely the case for 'now' and perhaps 'I', but it is at least possible in principle. After thinking (27), upon moving away from the place while keeping track of it, it would be possible to think:

(30) I was there a moment ago

The terms 'I', 'there' and 'a moment ago' could express the same singular modes of presentation as were expressed by 'I', 'here' and 'now' in (27). Furthermore if we were to accept that egocentric predicates are not strictly part of the thought expressed (though in section 6.5 it was suggested that this might not be the case) then it would follow that (30) could express the very same thought as (27). Yet (30) is by no means knowable a priori.

In fact if the references of the indexical terms are identified demonstratively then even (27) is not strictly knowable a priori. For in principle it may be false. Suppose that the reference of 'now' is identified in virtue of being the time at which a star is seen to explode. Given that this would actually have occurred in the distant past the thought

²³ I shall consider only cases in which the identifications are all demonstrative or all egocentric; it will be clear from this how cases in which there is a mixture of the two can be treated.

expressed by (27) would be false. (27) is only guaranteed to be true when there is no possibility of the indexical terms being used inappropriately, and this was shown in the previous section never to be entirely the case when there is a demonstrative identification of the reference.

This error would not be likely to occur in practice. As I have said, I do not suggest that the references of 'I' and 'now' are normally identified demonstratively (I suggest only that there are circumstances in which they could be identified in this way). And although I do suggest that demonstrative identifications of 'here' are common it does not appear that there are many cases in which I could be mistaken in thinking that I am located within the region which I refer to as 'here'. In cases of refraction like the one described above, for example, although the region referred to as 'here' is not co-extensive with the region which has the egocentric property 'here' I am nonetheless located within the relevant region (within both regions, in fact). I cannot think of a refractive illusion in which this would not be the case, though such an illusion might be possible. The only kind of spatial illusion I can think of in which (27) would come out false would be the 'virtual reality' kind. I could, for example, be unknowingly wearing virtual reality spectacles and could think that 'I am here now' where 'here' referred demonstratively to the place which I perceived. It might, however, be necessary to rule out cases like this; firstly for Evans's reason that the information is not received in the 'normal' way and secondly because it is not clear that it is possible to demonstratively identify a region of space in this way at all. It could very well be argued that a 'conceptual element' would be needed in such an identification, just as it is when one identifies a place seen on a television screen. If one does not know that one is looking at a screen one does not identify the place.

If illusions could be ruled out, however, then even with a demonstrative identification one could, after all, know (27) a priori. This is because of the relations between the egocentric predications. This becomes clearer in the case in which 'I', 'here' and 'now' are all identified egocentrically. Egocentric identification makes the

egocentric predications ‘is me’, ‘is here’ and ‘is now’ immune to error because, as described above, the reference is identified by virtue of satisfying those predicates. Now, if we assume once again (just for simplicity) that the egocentric predicates are not part of the thought expressed by (27) then the thought expressed by (27) is of the form:

(31) *S* is at place *p* at time *t*

This would be the same thought that is expressed by (30).²⁴ As with Kaplan’s account there is something common to (27) and (30) and it is this common element which is contingent; but on this account what is common to (27) and (30) is a Fregean thought rather than a *kontent*. Where (27) and (30) differ is in the associated egocentric predication rather than in their Kaplanian character; and it is this which makes (27), but not (30), knowable a priori.

The egocentric predication makes (27) knowable a priori because of the relations which a thinker ordinarily knows to hold between the egocentric properties ‘me’, ‘here’ and ‘now’. In particular, a thinker ordinarily knows that if a place is ‘here’ then it is ‘where I am now’ (the ‘I’ and ‘now’ imply the egocentric predications ‘is me’ and ‘is now’). It is not obvious that someone who grasped only the monadic form of the predicate ‘is here’ would know this but I assume that ordinary thinkers (anyone likely to reflect on the nature of (27)) grasp the relational predicate as well. Consequently someone who thinks (27) can infer that it is true if and only if:

(32) I am now where I am now

²⁴ If we drop the assumption that the egocentric predicate is not part of the thought expressed then (27) and (30) express different thoughts. The account would become slightly more complicated, but I hope it will be clear that these complications do not invalidate the account.

Clearly this is always true and its form makes it knowable a priori.

Incidentally it is worth noting that the terms 'I', 'here' and 'now' are not inter-definable. Although no one would suggest that understanding each of these terms consists entirely in understanding its relation to the other two (this would obviously be circular) there may nevertheless be a temptation to assume that given the reference of any two of these terms it is always possible to determine the reference of the third. This is not true. It is true that 'here' always refers to 'the place where I am now', and I have appealed to this in explaining how (27) can be known a priori. This one relation between 'I', 'here' and 'now' is sufficient for that purpose. But it is not true, for example, that the reference of 'now' is 'the time at which I am here' (because I may be here at many different times) and neither is it true that the reference of 'I' is 'the person who is now here'. The latter is because it is possible (logically, even if not in the actual world) for more than one person to be at a single place.²⁵ If I believed that I lived in a world where more than one person could occupy a place then I would not imagine that I could identify myself as 'the person who is now here'. Yet intuitively there seems to be no reason to imagine that my indexical thinking would be any different in such a world to the way it is in the actual world.

6.8 *Rip van Winkle*

I shall finish by briefly discussing the example with which Kaplan began the recent discussions of cognitive dynamics:

Suppose that yesterday you said, and believed it, 'It is a nice day today'. What does it mean to say, today, that you have retained *that* belief? It seems unsatisfactory to just believe the same content under any old character - where is the *retention*? You *can't* believe that content

under the same character. Is there some obvious standard adjustment to make to the character, for example, replacing *today* with *yesterday*? If so, then a person like Rip van Winkle, who loses track of time, can't retain any such beliefs. This seems strange (1989, pp. 537-8).

Rip van Winkle falls asleep and wakes up years later thinking that only a day has passed. Kaplan's theory implies that Rip fails to retain his belief because when he thinks 'yesterday was fine' the wrong day is referred to (because the character of 'yesterday' picks out the previous day, not the day on which Rip fell asleep). Kaplan expresses this as a worry; Evans (1985, p. 311) dismisses the worry, arguing that there is nothing more strange about the idea of someone losing a belief by losing track of time than there is in the idea that a belief about an object could be lost by losing track of the object.

John Perry (1997), who adopts Kaplan's content/character distinction at the level of both language and thought (he speaks of 'doxastic characters' for the latter), argues that Kaplan and Evans are both wrong to think that Rip cannot retain his belief:

When he awakes... Rip updates his belief according to his view of how the context has changed. His view about the change of context is mistaken, and the new character, 'Yesterday [the day before the day of this thought] was nice' is not a way of believing the original content. But that is no reason to say that Rip has lost his original belief. He retains it under various other backup characters (1997, p. 36).

Among the 'backup characters' which Rip might use to retain the belief, Perry lists '[the day I remember]' and '[the day this belief was acquired]'. Now, although I do not accept the content/character distinction at the level of thought because of the difficulties it faces in dealing with cognitive dynamics Perry's view has certain things in

²⁵ See Appendix B on the possibility of interpenetration of physical objects.

common with my own. The major difference is that Perry thinks that Rip can only retain the content of his belief; the fact that it is retained under a different character makes it a different psychological entity. I, by contrast, hold that Rip could, in principle, retain the very same Fregean thought. Whereas, for Perry, the retention of content requires backup characters on my view, the retention of a thought requires backup Ways of Referring. But Ways of Referring are not strictly part of the thought, even though in some cases they are part of the subject's overall psychological state.

Suppose that on the day on which he fell asleep Rip was able to identify the day not only egocentrically but also as, for example, the day on which a tornado was sighted near the town (assume that tornadoes are very rare in Rip's part of the world, so that this fact picks out a unique day during Rip's lifetime). When Rip wakes up he may assume that a day has passed and may therefore assume that the day he remembers can be identified as 'the day before today'. This is, of course, an error, just as Perry says. So long as Rip persists in this belief then he is surely in a confused state in which there is no unique day that he is thinking about when he says 'yesterday was nice'. But suppose he discovers that he has been asleep for longer than a day. Now his thought can become clear again, because he can identify the day as 'the day on which a tornado was sighted near the town' and dispense with the idea that it is also 'the day before today'. I see no reason why we cannot regard this as the retention of Rip's original thought. If no alternative Ways of Referring are available, however, Rip really does lose track of the thought.²⁶

²⁶ Some of Perry's examples of backup characters strike me as unable to satisfy Russell's Principle (just knowing that a day is the day I remember does not identify it adequately) but they are clearly supposed to perform the same identifying function as Ways of Referring. I assume he adopts these examples because they would always be available to Rip; whereas I, by contrast, think that there is only a possibility, rather than a guarantee, that Rip will be able to retain the thought using a backup Way of Referring.

APPENDIX C

An Argument for the Tenseless Theory of Time

I have argued that thoughts expressed using temporal indexicals involve or manifest an egocentric predication such as 'is now' or 'is past'. This is predicated of the time being thought about. It is possible, however, to dispute whether these predicates are really egocentric. The A-theory or 'tensed' theory of time says that the A-series of McTaggart's (1908) famous argument is real. This means that the properties picked out by terms like past, present and future, which are known in the literature as 'tenses', are mind-independent properties of times (and are therefore not egocentric). Consequently on this view the 'flow' or 'passage' of time is a feature of mind-independent reality. The B-theory or 'tenseless' theory says the opposite, that tenses have no mind-independent reality and that the apparent flow of time is merely psychological. Only the B-theory is consistent with the theory that tensed properties are egocentric.

Debates between A- and B-theorists have continued with increasing intensity in recent years.¹ Many of these debates have concerned the nature of temporal indexicals (and, in particular, whether they can be construed in a way which avoids McTaggart's paradox). Since these debates have become quite complex it would be useful to have an argument which settled the question about tense independently of questions about the functioning of indexical terms. This would settle the question as to whether the temporal predicates which I claim to be associated with temporal indexicals are egocentric without the need to become embroiled in further debates about tensed thought. It would not be a problem if the A-theory were correct; the theory of

indexical thoughts could simply be adjusted so that in the temporal case the predicates are not egocentric. Fortunately, however, I have an argument against the A-theory which does not depend on assumptions about temporal indexicals. I shall go through it in outline then explain it in more detail.

The argument, in a nutshell, is as follows. Experience provides each of us with a strong intuitive impression that time flows. The A-theorist takes this as a reflection of the fact that time really does flow. But on any physicalist or supervenience theory of the mind the nature of experience is determined entirely by the physical state of the world - experience could not be different unless the physical state of the world were different. Now our understanding of the role of time in physical science suggests that the putative flow of time has no role in determining the physical state of the world. It follows from this that the flow of time could have no role in determining the nature of experience. The impression that time flows thus arises quite independently of the putative real flow of time. Consequently the nature of temporal experience provides no reason to posit a real flow of time; and I shall argue that it follows in turn that there can be no a priori reason either. I shall now go through each step of this argument in detail.

It appears to be common ground among most A- and B-theorists that experience provides a strong intuitive impression that time flows. This is often assumed to favour the A-theory; for it seems to be implicitly assumed that the A-theory can account for the way time seems to flow by appealing to the supposed fact that it really does flow. A certain onus is placed instead on the B-theorists to account for the temporal features of experience in a way which is consistent with their denial that the flow of time is real. Thus L. Nathan Oaklander remarks that:

¹ For book-length defences of the A- and B-theories see Smith 1993 and Mellor 1998 respectively. For further relevant examples of current debates see the papers collected in Oaklander and Smith 1994 and also Le Poidevin 1998.

Regardless of how many arguments Mellor, Smart, I, and other defenders of the tenseless theory offer, we will never persuade tenses to abandon their view unless we can explain those features of our experience of time that seem to require that time be tensed. It is an impression deeply felt by all of us that time, or events in time, flow from the far future to the near future and then, after shining in the spotlight of the NOW or present moment, immediately recede into the more and more distant past. But does this impression reflect a basic truth about time? (Oaklander and Smith 1994, p. 289).

Accounting for the temporal features of experience certainly poses a difficult and important task for the B-theory. But the A-theory owes us an account of these features of experience too - one which shows how the putative real flow of time can be a factor in determining the nature of temporal experience. I shall argue that no such account is possible.

The reason why the A-theorist's task is impossible is that the flow of time, as it is normally construed, couldn't cause the kinds of things it would have to cause in order for it to have a role in explaining why temporal experience is the way it is. The features of experience that give the impression that time flows are part of the way things seem to the subject; they are not merely relational states. That is to say, the kind of temporal experience that we have could not consist in merely experiencing times which are as a matter of fact tensed because the putative tenses show up in the way things seem to the subject. Now on any physicalist or supervenience theory of the mind the way things seem to the subject is determined by some aspect of the brain state; consequently if there are such things as mind-independent tenses they must have a role in determining brain states.

However only certain kinds of things can have a role in determining brain states. The brain is a part of the physical world and the succession of its states is determined by the physics of the brain and its environment. If someone were isolated within a closed environment then in principle the physics of the system comprising the brain and

its environment would determine the subsequent states of the brain and thus the experiential states of the subject. Even if a physicist could not predict these states in practice, and even allowing for quantum indeterminacy and the like, the system would still evolve in accordance with the laws of physics and nothing else.

Now the laws of physics do not require that time flows. The mathematical equations which describe the evolution of a system over time contain a variable t for time, into which one can substitute arbitrary values to determine the state of the system at a given time. The notion of flow does not enter the picture at all. Physics does of course require there to be a temporal ordering (a B-series) and the position of a given event in this ordering can be relevant in determining its effects. But this position is relevant only because it is relative to other events; the absolute temporal locations determined by the tensed properties of an A-series have no such role in determining physical states.

Although in the past some writers have thought that the tenseless nature of physical laws was alone sufficient to refute the A-theory, that is not the argument to be given here.² Moreover it is of course necessary to use tensed vocabulary when applying physical laws in practice; there will be statements like 'it is now three o'clock and the system is in state S_1 ; the laws of physics show that the system will be in state S_2 five minutes after it is in state S_1 , so it will be in state S_2 five minutes from now'. But what matters for the present argument is just that no tensed properties have a role in determining the states S_1 and S_2 themselves; they are the same physical states no matter whether they occur in the past, present or future.

These considerations make it clear that if there were such a thing as the flow of time it would have no role in determining brain states and consequently it would have no role in determining the nature of experience. The impression that time flows thus arises quite independently of whether or not it really flows. This is not to suggest that the putative flow of time must be construed as altogether impotent. All that need be said is that the flow of time would be powerless to cause the sorts of things it would have to

cause in order to be experienced. Or, to put it in a way which avoids mentioning causation at all, the flow of time can have no role in determining the nature of temporal experience because it plays no role in determining brain states. The argument does not say that we should reject the existence of anything which cannot exert causal influence on us; it applies only to what is reflected in experience.

There is a useful analogy between the present argument and a certain argument against those theories in the philosophy of mind which posit epiphenomenal qualia. Epiphenomenal qualia are supposed to be features of consciousness which give our experiences the particular subjective qualities they have. But although the putative epiphenomenal qualia are effects - they are caused by states of the brain - they themselves have no effects in the physical world whatever, and consequently are missed out by physical descriptions of the world.

Daniel Dennett points out what is wrong with this view in *Consciousness Explained*:³

Suppose that Otto insists that he (for one) has epiphenomenal qualia. Why does he say this? Not because they have some effect on him, somehow guiding him or alerting him as he makes his avowals. By the very definition of epiphenomena (in the philosophical sense), Otto's heartfelt avowals that he has epiphenomena could not be evidence for himself or anyone else that he does have them, since he would be saying exactly the same thing even if he didn't have them (Dennett 1991, pp. 402-3).

Consequently we cannot have any evidence for the existence of epiphenomenal qualia, and there seems no sensible option but to reject them. I suggest that we can regard the flow of time as similarly epiphenomenal. Our avowals that time seems to flow can be no evidence that it does flow, for we would be saying the same things even if it did not.

² See for instance Russell 1940 and Grünbaum 1967.

³ Dennett's argument follows those of I. Fox (1989) and G. Harman (1990).

I have shown that experience would be just the way it is even on the assumption that there is no real flow of time. Since experience thus provides no support for the A-theory, it seems in order to reject it. But perhaps the A-theorist will now claim that there are arguments in favour of the A-theory which do not appeal to features of experience, and which are thus immune to the above criticisms. If the above considerations are accepted, however, any such a priori arguments would face a formidable difficulty. The problem would be to make clear just what was being claimed. Given the above arguments, the a priori A-theorist would have to accept that whatever the putative A-properties are, they have nothing to do with what is reflected in experience. It would therefore be necessary to find suitable terms with which to refer to these putative A-properties. But I took as a starting point the fact that experience provides a strong impression that time flows. We can of course report on this using tensed vocabulary (as Oaklander does when he uses the words 'past', 'present' and 'future' in the quotation above). Consequently the normal tensed vocabulary is unavailable to the A-theorist because it is already in use to describe what is reflected in experience, which it has already been accepted has nothing to do with A-properties. This implies that any intrinsic properties of times posited by an a priori argument can have nothing to do with pastness, presentness or futurity. So nothing recognisable as the A-theory can be salvaged.

Carnap and others used to reject as 'metaphysics' anything which could not be reduced to experience. It is important to see that the present argument, by contrast, concerns only that which is reflected in experience. It shows that given a physicalist or supervenience theory of the mind nothing that is reflected in experience can belong within the realm of metaphysics. The dualism of the A- and B-series, like the Cartesian dualism of mental and physical substance, comes unstuck through a combination of the impossibility of explaining how the two sides could interact and the unintelligibility of claiming that they do not. A modern understanding of the place of the mind in the

natural world thus shows the flow of time to be nothing more than a last imaginary ghost in nature's machinery.

Bibliography

- Ackerman, Diana 1979: 'Proper Names, Propositional Attitudes and Non-Descriptive Connotations'. *Philosophical Studies*, 35, pp. 55-69.
- Almog, J., Perry, J. and Wettstein, H. (eds.) 1989: *Themes from Kaplan*. Oxford: Oxford University Press.
- Anscombe, G. E. M. 1975: 'The First Person', in Guttenplan, Samuel (ed.) *Mind and Language: Wolfson College Lectures, 1974*. Oxford: Oxford University Press. Reprinted in Cassam 1994.
- Bach, K. 1994: *Thought and Reference*. Oxford: Clarendon Press.
- Beaney, Michael, (ed.) 1997: *The Frege Reader*. Oxford: Blackwell.
- Bermúdez, José Luis 1997: 'Practical Understanding vs. Reflective Understanding'. *Philosophy and Phenomenological Research*, 57, pp. 635-641.
- 1998: *The Paradox of Self-Consciousness*. Cambridge MA/London: MIT Press/A Bradford Book.
- Bezuidenhout, Anne 1996: 'Pragmatics and Singular Reference'. *Mind and Language*, 11, pp. 133-159.
- Bilgrami, Akeel 1992: *Belief and Meaning*. Oxford: Blackwell.
- Blackburn, Simon 1984: *Spreading the Word*. Oxford: Clarendon Press.
- Brandom, Robert 1994: *Making it Explicit*. Cambridge, MA/London: Harvard University Press.
- 1997: 'Précis of *Making it Explicit*', *Philosophy and Phenomenological Research*, 57, pp. 153-156.
- Brewer, Bill 1999: *Perception and Reason*. Oxford: Clarendon Press.
- Burge, Tyler 1979: 'Individualism and the Mental'. *Midwest Studies in Philosophy*, 4, pp. 73-121.
- 1982: 'Other Bodies', in Woodfield, A. (ed.) *Thought and Object*. Oxford: Oxford University Press.
- 1982a, 'Two Thought Experiments Reviewed'. *Notre Dame Journal of Formal Logic*, 23, pp. 284-293.
- 1994: 'Individualism and Self-Knowledge'. *Journal of Philosophy*, 85, pp. 649-663. Reprinted in Cassam 1994.
- Campbell, John 1987: 'Is Sense Transparent?'. *Proceedings of the Aristotelian Society*, 88, pp. 273-292.
- 1993: 'A Simple View of Colour', in Haldane, John and Wright, Crispin (eds.) *Reality, Representation and Projection*. Oxford: Oxford University Press.
- 1993a: 'The Role of Physical Objects in Spatial Thinking', in Eilan, McCarthy and Brewer (eds.) 1993, pp. 65-95.
- 1994: *Past, Space and Self*. Cambridge, MA/London: MIT Press/A Bradford Book.
- 1997: 'Précis of *Past, Space and Self*'. *Philosophy and Phenomenological Research*, 57, pp. 633-634.
- 1997a: 'Replies'. *Philosophy and Phenomenological Research*, 57, pp. 655-670.

- 1998: 'Joint Attention and the First Person', in O'Hear, Anthony (ed.) *Current Issues in Philosophy of Mind: Royal Institute of Philosophy Supplement 43*. Cambridge: Cambridge University Press, pp. 123-136.
- 2001: 'Memory Demonstratives', in Hoerl and McCormack (eds.) 2001, pp. 169-186.
- Cappelen, Herman 1999: 'Intentions in Words'. *Noûs*, 33, pp. 92-102.
- Carruthers, Peter 1984: 'Fragmentary Sense'. *Mind*, 93, pp. 351-369.
- 1988: 'More Faith than Hope: Russellian Thoughts Attacked'. *Analysis*, 48, pp. 91-96.
- Cassam, Quassim (ed.) 1994: *Self-Knowledge*. Oxford: Oxford University Press.
- 1997: 'Subjects and Objects'. *Philosophy and Phenomenological Research*, 57, pp. 643-648.
- Castañeda, Hector-Neri 1969: 'On the Phenomeno-Logic of the I'. *Proceedings of the 14th International Congress of Philosophy, iii*. Vienna: University of Vienna, pp. 260-266. Reprinted in Cassam 1994.
- Child, William 1994: *Causality, Interpretation and the Mind*. Oxford: Clarendon Press.
- Christofidou, Andrea 1995: 'First Person: The Demand for Identification-Free Self-Reference'. *The Journal of Philosophy*, 92, pp. 223-234.
- 2000: 'Self-Consciousness and the Double Immunity'. *Philosophy*, 75, pp. 539-569.
- Cockburn, David 1997: *Other Times*. Cambridge: Cambridge University Press.
- Coates, Paul 2000: 'Deviant Causal Chains and Hallucinations: a Problem for the Anti-Causalist'. *The Philosophical Quarterly*, 50, pp. 320-331.
- Collins, Arthur 1997: 'The Psychological Reality of Reasons'. *Ratio*, 10, pp. 108-123.
- Corazza, Eros, and Jérôme Dokic 1992: 'On the Cognitive Significance of Indexicals'. *Philosophical Studies*, 66, pp. 183-196.
- Craig, William Lane 1998: 'McTaggart's Paradox and the Problem of Temporary Intrinsic'. *Analysis*, 58, pp. 122-127.
- 2000: 'Why is it Now?'. *Ratio*, 13, pp. 115-122.
- 2001: 'Wishing It Were Some Other Time'. *Philosophy and Phenomenological Research*, 62, pp. 159-166.
- Crawford, Sean 1998: 'In Defence of Object-Dependent Thoughts'. *Proceedings of the Aristotelian Society*, 98, pp. 201-210.
- Dancy, Jonathan 2000: *Practical Reality*. Oxford: Oxford University Press.
- Davidson, Donald 1976: 'Reply to Foster', in Evans, Gareth and McDowell, John (eds.) *Truth and Meaning: Essays in Semantics*. Oxford: Oxford University Press. Reprinted in Davidson 1984.
- 1984: *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- 1984a: 'First Person Authority'. *Dialectica*, 38, pp. 101-111.
- 1986: 'A Nice Derangement of Epitaphs', in LePore, Ernest (ed.) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Blackwell.
- 1987: 'Knowing One's Own Mind'. *The Proceedings and Addresses of the American Philosophical Association*, 60, pp. 441-58. Reprinted in Cassam 1994.
- 1990: 'A Coherence Theory of Truth and Knowledge' (including 'Afterthoughts, 1987'); in Malachowski, Alan (ed.) *Reading Rorty*. Oxford: Blackwell.
- 1990a: 'Meaning, Truth and Evidence', in Barrett, R. B. and Gibson, R. F. (eds.) *Perspectives on Quine*. Oxford: Blackwell.

- 1991: 'Epistemology Externalised'. *Dialectica*, 45, pp. 191-202.
- 1989: 'What is Present to the Mind?', in Brandl, J. and Gombocz, W. (eds.) *The Mind of Donald Davidson (Grazer Phil. Stud. 36)*. Amsterdam: Rodopi. pp. 3-18. Reprinted in Villanueva 1991, pp. 197-213.
- 1996: 'The Folly of Trying to Define Truth'. *The Journal of Philosophy*, 93, pp. 263-278.
- DeBellis, Mark 1995: *Music and Conceptualisation*. Cambridge: Cambridge University Press.
- Dennett, Daniel C. 1975: 'Why the Law of Effect Will Not Go Away'. *Journal of the Theory of Social Behaviour*, 5, pp. 169-187. Reprinted in *Brainstorms: Philosophical Essays on Mind and Psychology*. Montgomery: Bradford Books, 1978.
- 1987: *The Intentional Stance*. Cambridge, MA/London: MIT Press/A Bradford Book.
- 1988: 'Quining Qualia', in Marcel, A. J. and Bisiach, E. (eds.) *Consciousness in Contemporary Science*. Oxford: Clarendon Press.
- 1991: *Consciousness Explained*. London: Penguin.
- Dokic, Jérôme 1996: 'The Dynamics of Deictic Thoughts'. *Philosophical Studies*, 82, pp. 179-204.
- Donnellan, Keith 1966: 'Reference and Definite Descriptions'. *Philosophical Review*, 75, pp. 281-304.
- Dretske, Fred I. 1981: *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press. Reprinted Stanford: CSLI Publications, 1999.
- 1989: 'Reasons and Causes', in Tomberlin, James E. (ed.) *Philosophical Perspectives 3*. Atascadero, CA: Ridgeview.
- 1993: 'Conscious Experience'. *Mind*, 102, pp. 263-283.
- Dummett, Michael 1960: 'A Defence of McTaggart's Proof of the Unreality of Time'. *Philosophical Review*, 69, pp. 497-504. Reprinted in Dummett 1978.
- 1975: 'What is a Theory of Meaning? (I)' in Guttenplan, Samuel (ed.) *Mind and Language*. Oxford: Oxford University Press. Reprinted in Dummett 1993.
- 1976: 'What is a Theory of Meaning? (II)' in Evans, Gareth and McDowell, John (eds.) *Truth and Meaning: Essays in Semantics*. Oxford: Clarendon Press. Reprinted in Dummett 1993.
- 1978: *Truth and Other Enigmas*. London: Duckworth.
- 1978a: 'Frege's Distinction Between Sense and Reference', in Dummett 1978.
- 1978b: 'Frege's Philosophy', in Dummett 1978.
- 1978c: 'The Reality of the Past', in Dummett 1978.
- 1981: *Frege: Philosophy of Language*. 2nd edition. London: Duckworth.
- 1981a: *The Interpretation of Frege's Philosophy*. London: Duckworth.
- 1991: *Frege: Philosophy of Mathematics*. London: Duckworth.
- 1993: *The Seas of Language*. Oxford: Clarendon Press.
- 1993a: *Origins of Analytical Philosophy*. London: Duckworth.
- 1996: *Frege and Other Philosophers*. Oxford: Clarendon Press.
- Durrant, Michael 1988: 'Demonstrative Identification and Descriptive Identification'. *Philosophical Quarterly*, 38, pp. 213-221.
- Eilan, N., McCarthy, R. and Brewer, B. (eds.) 1993: *Spatial Representation: Problems in Philosophy and Psychology*. Oxford: Blackwell.

- Eilan, Naomi 1995: 'The First Person Perspective'. *Proceedings of the Aristotelian Society*, 95, pp. 51-66.
- Evans, Gareth 1973: 'The Causal Theory of Names'. *Aristotelian Society Supplementary Volume*, 47, pp. 187-208. Reprinted in Evans 1985.
- 1975: 'Identity and Predication'. *The Journal of Philosophy*, 72, pp. 343-363. Reprinted in Evans 1985.
- 1980: 'Things Without the Mind - A Commentary upon Chapter Two of Strawson's *Individuals*', in Van Straaten, Zak (ed.) *Philosophical Subjects: Essays Presented to P. F. Strawson*. Oxford: Clarendon Press. pp. 76-116.
- 1981: 'Understanding Demonstratives', in Parret, Herman and Bouveresse, Jacques (eds.) *Meaning and Understanding*. Berlin/New York: De Gruyter. pp. 280-303. Reprinted in Evans 1985.
- 1982: *The Varieties of Reference*. Edited by John McDowell. Oxford: Clarendon Press.
- 1985: *Collected Papers*. Oxford: Clarendon Press.
- 1985a: 'Molyneux's Question', in Evans 1985.
- Ezcurdia, Maite 1997: 'Dynamic and Coherent Thoughts', in Dokic, Jérôme (ed.) *European Review of Philosophy*, vol. 2: *Cognitive Dynamics*. Stanford, CA: CSLI Publications. pp. 105-139.
- Farrell Smith, Janet 1989: 'Russell on Indexicals and Scientific Knowledge', in Wade Savage, C. and Anderson, C. Anthony (eds.) *Minnesota Studies in the Philosophy of Science*, vol. 12. Minneapolis: University of Minnesota Press. pp. 119-137.
- Fodor, Jerry A. 1975: *The Language of Thought*. New York: Thomas Y. Crowell.
- 1980: 'Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology'. *The Behavioural and Brain Sciences*, 3, pp. 63-109.
- 1983: *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge MA/London: MIT Press/A Bradford Book.
- 1987: *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA/London: MIT Press/A Bradford Book.
- 1989: 'Substitution Arguments and the Individuation of Beliefs', in Boolos, G. (ed.) *Method, Reason and Language*. Cambridge: Cambridge University Press. Reprinted in Fodor 1990.
- 1990: *A Theory of Content and Other Essays*. Cambridge, MA/London: MIT Press/A Bradford Book.
- 1994: *The Elm and the Expert: Mentalese and its Semantics*. Cambridge, MA/London: MIT Press/A Bradford Book.
- 1998: *Concepts: Where Cognitive Science Went Wrong*. Oxford: Clarendon Press.
- 1998a: 'There Are No Recognitional Concepts - Not Even RED (Parts 1 & 2)', in Fodor's *In Critical Condition: Polemical Essays on Cognitive Science and the Philosophy of Mind*. Cambridge, MA/London: MIT Press/A Bradford Book.
- 2000: 'Replies to Critics'. *Mind & Language*, 15, pp. 350-374.
- Fox, I. 1989: 'On the Nature and Cognitive Function of Phenomenal Content - Part One'. *Philosophical Topics*, 17, pp. 81-117.
- van Fraassen, Bas C. 1966: 'Singular Terms, Truth-Value Gaps, and Free Logic'. *Journal of Philosophy*, 63, pp. 481-495.
- Frege, Gottlob 1892: 'Über Sinn und Bedeutung'. *Zeitschrift für Philosophie und Philosophische Kritik*, 100, pp. 25-50. Translated as Frege 1966.

- 1956: 'The Thought: A Logical Inquiry'. Translated by A. M. and Marcelle Quinton. *Mind*, 65, pp. 289-311.
- 1966: 'On Sense and Reference' in *Translations from the Philosophical Writings of Gottlob Frege*. Translated by P. Geach and M. Black. 3rd edn. Oxford: Blackwell. pp. 56-78 (translation of Frege 1892).
- 1978: *Die Grundlagen der Arithmetik*, translated by J. L. Austin as *The Foundations of Arithmetic*. 2nd edn. (revised). Oxford: Blackwell.
- 1979: *Posthumous Writings*. Edited by H. Hermes, F. Kambartel and F. Kaulbach, translated by P. Long and R. White. Oxford: Blackwell.
- 1980: *Philosophical and Mathematical Correspondence*. Translated by Hans Kaal. Oxford: Blackwell.
- 1980a: 'On Concept and Object', in *Translations from the Philosophical Writings of Gottlob Frege*. Translated by P. Geach and M. Black. 3rd edn. Oxford: Blackwell. pp. 42-55.
- 1997: 'Logic', in Beaney 1997, pp. 227-250.
- Friedman, William 1990: *About Time: Inventing the Fourth Dimension*. Cambridge, MA: MIT Press/A Bradford Book.
- Gallois, André 1998: *Occasions of Identity: A Study in the Metaphysics of Persistence, Change, and Sameness*. Oxford: Clarendon Press.
- García-Carpintero, Manuel 1998: 'Indexicals as Token-Reflexives'. *Mind*, 107, pp. 529-563.
- Geach, P. T. 1968: *A History of the Corruptions of Logic*. University of Leeds Inaugural Lecture. Leeds: Leeds University Press.
- Goodman, Nelson 1965: *Fact, fiction and forecast*. 2nd ed. Indianapolis: Bobbs-Merrill.
- Gregory, Richard L. 1970: *The Intelligent Eye*. London: Weidenfeld & Nicolson.
- 1981: *Mind in science: A History of Explanations in Psychology and Physics*. London: Weidenfeld and Nicolson.
- (ed.) 1987: *The Oxford Companion to the Mind*. Oxford: Oxford University Press.
- Grice, H. P., and Strawson, P. F. 1956: 'In Defense of a Dogma'. *Philosophical Review*, 65, pp. 141-158.
- Grice, H. P. 1957: 'Meaning'. *Philosophical Review*, 66, pp. 377-388.
- 1969: 'Vacuous Names', in Davidson, Donald and Hintikka, Jaakko (eds.) *Words and Objections: Essays on the Work of W. V. Quine*. Dordrecht: Reidel. pp. 118-145.
- Grimaltos, Tobies, and Moya, Carlos J. 1997: 'Belief, Content and Cause', in Dokic, Jérôme (ed.) *European Review of Philosophy, vol. 2: Cognitive Dynamics*. Stanford, CA: CSLI Publications. pp. 159-171.
- Grünbaum, Adolf 1967: *Modern Science and Zeno's Paradoxes*. Middletown, CT: Wesleyan University Press.
- Hale, Bob, and Wright, Crispin (eds.) 1997: *A Companion to the Philosophy of Language*. Oxford: Blackwell.
- Harman, G. 1990: 'The Intrinsic Quality of Experience', in Tomberlin, J. E. (ed.) *Philosophical Perspectives, 4: Action Theory and Philosophy of Mind*. Atascadero, CA: Ridgeview.
- Heck, Richard G. Jr. 1995: 'The Sense of Communication'. *Mind*, 104, pp. 79-106.
- van Heijenoort, J. 1977: 'Frege on Sense Identity'. *Journal of Philosophical Logic*, 6, pp. 103-108.

- Hestevold, H. Scott 1990: 'Passage and the Presence of Experience'. *Philosophy and Phenomenological Research*, 50, pp. 537-552. Reprinted in Oaklander and Smith 1994.
- Higginbotham, James 1995: 'Tensed Thoughts'. *Mind & Language*. 10, pp. 226-249.
- Hoerl, Christoph 1997: 'Cognitive Dynamics: An Attempt at Changing Your Mind', in Dokic, Jérôme (ed.) *European Review of Philosophy*, vol. 2: *Cognitive Dynamics*. Stanford, CA: CSLI Publications. pp. 141-158.
- 1998: 'The Perception of Time and the Notion of a Point of View'. *European Journal of Philosophy*. 6, pp. 156-171.
- and McCormack, Teresa (eds.) 2001: *Time and Memory: Issues in Philosophy and Psychology*. Oxford: Clarendon Press.
- Hookway, Christopher 1988: *Quine*. Cambridge: Polity Press.
- Hume, David 1978: *A Treatise of Human Nature*. Edited by L. A. Selby-Bigge, 2nd edn., revised by P. H. Nidditch. Oxford: Clarendon Press.
- Hylton, Peter 1989: 'The Significance of "On Denoting"', in Savage, C. Wade and Anderson, C. Anthony (eds.) *Minnesota Studies in the Philosophy of Science*, vol. 12. Minneapolis: University of Minnesota Press. pp. 88-107.
- Johnston, Mark 1987: 'Is There a Problem about Persistence?'. *Aristotelian Society Supplementary Volume*, 61, pp. 107-135.
- Kant, Immanuel 1933: *Critique of Pure Reason*. Translated by Norman Kemp Smith. London: Macmillan.
- Kaplan, David 1978: 'Dthat', in Cole, P. (ed.) *Syntax and Semantics*. New York: Academic Press. pp. 221-243. Reprinted in Yourgrau 1990.
- 1989: 'Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and other Indexicals', in Almog, Perry and Wettstein (eds.) 1989.
- 1989a: 'Afterthoughts', in Almog, Perry and Wettstein (eds.) 1989.
- 1990: 'Thoughts on Demonstratives', in Yourgrau (ed.) 1990. pp. 34-49.
- 1990a: 'Words'. *Aristotelian Society Supplementary Volume*, 64, pp. 93-119.
- Kirk, Robert 1994: 'Review of Christopher Peacocke's *A Study of Concepts*'. *Philosophical Books*, 35, pp. 51-54.
- Kripke, Saul A. 1979: 'A Puzzle About Belief', in Margalit, A. (ed.) *Meaning and Use*. Dordrecht: D. Reidel.
- 1980: *Naming and Necessity*. Oxford: Blackwell. Originally published in G. Harman and D. Davidson (eds.), *Semantics of Natural Language*. Dordrecht: Reidel, 1972.
- 1982: *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press.
- Kuhn, Thomas S. 1962: *The Structure of Scientific Revolutions*. Chicago: Chicago University Press.
- Lemmon, E. J. 1987: *Beginning Logic*. 2nd edition. London: Chapman & Hall.
- Le Poidevin, Robin 1998: 'The Past, Present, and Future of the Debate about Tense', in Le Poidevin, Robin (ed.) *Questions of Time and Tense*. Oxford: Clarendon Press.
- 1999: 'Egocentric and Objective Time'. *Proceedings of the Aristotelian Society*, 99, pp. 19-36.
- Lepore, Ernest, and Ludwig, Kirk 2000: 'The Semantics and Pragmatics of Complex Demonstratives'. *Mind*, 109, pp. 199-240.
- Lewis, David 1979: 'Attitudes *De Dicto* and *De Se*'. *The Philosophical Review*, 88, pp. 513-543.

- 1980: 'Index, Context and Content', in Kanger, S. and Öhman, S. (eds.) *Philosophy and Grammar*. Dordrecht: Reidel.
- 1983: *Philosophical Papers, Volume 1*. Oxford: Oxford University Press.
- 1983a: 'Counterparts of Persons and Their Bodies', in Lewis 1983, pp. 47-54.
- 1986: *On the Plurality of Worlds*. Oxford: Blackwell.
- Luntley, Michael, 'Thinking of Individuals: a Prolegomenon to Any Future Theory of Thought', in *The Philosophy of P. F. Strawson*, ed. by Pranab Kumar Sen and Roop Rekha Verma (New Delhi: Indian Council for Philosophical Research, 1995).
- 1997: 'Dynamic Thoughts and Empty Minds', in Dokic, Jérôme (ed.) *European Review of Philosophy, vol. 2: Cognitive Dynamics*. Stanford, CA: CSLI Publications. pp. 77-103.
- 1998: 'The "Practical Turn" and the Convergence of Traditions'. *Philosophical Explorations*, 1.
- 1999: *Contemporary Philosophy of Thought: Truth, World, Content*. Oxford: Blackwell.
- Martin, Michael 1993: 'The Rational Role of Experience'. *Proceedings of the Aristotelian Society*, 93, pp. 71-88.
- Martinich, A. P., (ed.) 1996: *The Philosophy of Language*. Third edition. Oxford: Oxford University Press.
- McCulloch, Gregory 1988: 'Faith, Hope and Charity: Russellian Thoughts Defended'. *Analysis*, 48, pp. 84-90.
- 1991: 'Making Sense of Words'. *Analysis*, 51, pp. 73-79.
- McDowell, John 1977: 'On the Sense and Reference of a Proper Name'. *Mind*, 86, pp. 159-185.
- 1982: 'Truth-Value Gaps', in Cohen, L. J., Los, J., Pfeiffer, H. and Podewski, K. (eds.) *Logic, Methodology and Philosophy of Science VI*. Amsterdam/New York/Oxford: North-Holland.
- 1982a: 'Criteria, Defeasibility, and Knowledge'. *Proceedings of the British Academy*, 68, pp. 455-479.
- 1984: 'De Re Senses'. *The Philosophical Quarterly*, 34, pp. 283-294.
- 1985: 'Values and Secondary Qualities', in Honderich, Ted (ed.) *Morality and Objectivity: A Tribute to J. L. Mackie*. London: Routledge.
- 1986: 'Singular Thought and the Extent of Inner Space', in Pettit and McDowell (eds.) 1986, pp. 137-168.
- 1990: 'Peacocke and Evans on Demonstrative Content'. *Mind*, 99, pp. 255-266.
- 1991a: 'Intentionality De Re', in LePore, Ernest and van Gulick, Robert (eds.) *John Searle and His Critics*. Oxford: Blackwell. Reprinted in McDowell 1998.
- 1993: 'Knowledge by Hearsay', in Matilal, B. K. and Chakrabarti, A. (eds.) *Knowing from Words: Western and Indian Philosophical Analysis of Understanding and Testimony*. Dordrecht: Kluwer. Reprinted in McDowell 1998.
- 1994: *Mind and World*. Cambridge, MA/London: Harvard University Press.
- 1994a: 'The Content of Perceptual Experience'. *The Philosophical Quarterly*, 44, pp. 190-205.
- 1995: 'Knowledge and the Internal'. *Philosophy and Phenomenological Research*, 55, pp. 877-893.
- 1997: 'Brandom on Representation and Inference'. *Philosophy and Phenomenological Research*, 57, pp. 157-162.

- 1997a: 'Reductionism and the First Person', in Dancy, Jonathan (ed.) *Reading Parfit*, Oxford: Blackwell. pp. 230-250.
- 1998: *Meaning, Knowledge, and Reality*. Cambridge MA/London: Harvard University Press.
- 1998a: *Mind, Value, and Reality*. Cambridge MA/London: Harvard University Press.
- 1998b: 'Précis of *Mind and World*'. *Philosophy and Phenomenological Research*, 58, pp. 365-369.
- 1998c: 'Reply to Commentators'. *Philosophy and Phenomenological Research*, 58, pp. 381-388.
- McTaggart, J. M. E. 1908: 'The Unreality of Time'. *Mind*, 18, pp. 457-484.
- Melia, Joseph 2000: 'Continuants and Occurrents'. *Aristotelian Society Supplementary Volume*, 74, pp. 77-92.
- Mellor, D. H. 1981: *Real Time*. Cambridge: Cambridge University Press.
- Mellor, D. H. 1998: *Real Time II*. London/New York: Routledge.
- Miller, David (ed.) 1985: *Popper Selections*. Princeton, NJ: Princeton University Press.
- Miller, David 1994: *Critical Rationalism: A Restatement and Defence*. Chicago/La Salle, Illinois: Open Court.
- Millikan, Ruth Garrett 1993: *White Queen Psychology and Other Essays for Alice*. Cambridge MA/London: MIT Press/A Bradford Book.
- Newton, Natika 1992: 'Dennett on Intrinsic Intentionality'. *Analysis*, 52, pp. 18-23.
- Noonan, H. W. 1991: 'Object-dependent Thoughts and Psychological Redundancy'. *Analysis*, 51, pp. 1-9.
- 1993: 'Object-Dependent Thoughts: A Case of Superficial Necessity but Deep Contingency?', in Heil, J. and Mele, A. (eds.) *Mental Causation*. Oxford: Clarendon Press. pp. 283-308.
- Oaklander, L. Nathan 1993: 'On the Experience of Tenseless Time'. *Journal of Philosophical Research*, 18, pp. 159-166. Reprinted in Oaklander and Smith 1994.
- and Quentin Smith (eds.) 1994: *The New Theory of Time*. New Haven and London: Yale University Press.
- O'Brien, Lucy F. 1995: 'Evans on Self-Identification'. *Noûs*, 29, pp. 232-247.
- Oderberg, David S. 1993: *The Metaphysics of Identity Over Time*. London: The Macmillan Press.
- Papathomas, Thomas V. 1999: 'The Brain as a Hypothesis-constructing-and-testing Agent', in Lepore, Ernest and Pylyshyn, Zenon (eds.) *What is Cognitive Science?* Oxford: Blackwell. pp. 230-247.
- Peacocke, Christopher 1981: 'Demonstrative Thought and Psychological Explanation'. *Synthese*, 49, pp. 187-217.
- 1983: *Sense and Content*. Oxford: Clarendon Press.
- 1986: *Thoughts: An Essay on Content*. Oxford: Blackwell.
- 1986a: 'Analogue Content'. *Aristotelian Society Supplementary Volume*, 60, pp. 1-17.
- 1991: 'Demonstrative Content: A Reply to John McDowell'. *Mind*, 100, pp. 123-133.
- 1992: *A Study of Concepts*. Cambridge MA/London: MIT Press/A Bradford Book.
- 1993: 'Externalist Explanation'. *Proceedings of the Aristotelian Society*, 93, pp. 203-230.

- 1993b: ‘Intuitive Mechanics, Psychological Reality and the Idea of a Material Object’, in Eilan, McCarthy and Brewer 1993, pp. 162-176.
- 1997: ‘Holism’, in Hale and Wright 1997, pp. 227-247.
- 1998: ‘Nonconceptual Content Defended’. *Philosophy and Phenomenological Research*, 58, pp. 381-388.
- 2000: ‘Fodor on Concepts: Philosophical Aspects’. *Mind & Language*, 15, pp. 327-340.
- 2000a: ‘Theories of Concepts: A Wider Task’. *European Journal of Philosophy*, 8, pp. 298-321.
- Perry, John 1977: ‘Frege on Demonstratives’. *The Philosophical Review*, 86, pp. 474-497.
- 1979: ‘The problem of the Essential Indexical’. *Noûs*, 13, pp. 3-21.
- 1980: ‘A Problem about Continued Belief’. *Pacific Philosophical Quarterly*, 61, pp. 317-332. Reprinted in Perry 1993.
- 1988: ‘Cognitive Significance and New Theories of Reference’. *Noûs*, 22, pp. 1-18. Reprinted in Perry 1993.
- 1993: *The Problem of the Essential Indexical and Other Essays*. Oxford: Oxford University Press.
- 1993a: ‘Postscript to “Frege on Demonstratives”’, in Perry 1993.
- 1997: ‘Rip van Winkle and Other Characters’, in Dokic, Jérôme (ed.) *European Review of Philosophy*, vol. 2: *Cognitive Dynamics*. Stanford, CA: CSLI Publications. pp. 13-39.
- Persson, Ingmar 1999: ‘Awareness of One’s Body as Subject and Object’. *Philosophical Explorations*, 2, pp. 70-76.
- Pettit, Philip, and McDowell, John (eds.) 1986: *Subject, Thought and Context*. Oxford: Clarendon Press.
- Pettit, Philip, and McDowell, John 1986a: ‘Introduction’, in Pettit and McDowell 1986, pp. 1-15.
- Picardi, Eva 1993: ‘Note on Dummett and Frege on Sense-Identity’. *European Journal of Philosophy*, 1, pp. 69-80.
- Popper, Karl R. 1959: *The Logic of Scientific Discovery*. London: Routledge.
- 1972: *Objective Knowledge*. Oxford: Oxford University Press.
- 1975: ‘Evolutionary Epistemology’, in Harré, R. (ed.) *Problems of Scientific Revolution*. Oxford: Oxford University Press. Reprinted in Miller 1985.
- 1989: *Conjectures and Refutations*. Fifth edition. London: Routledge.
- Prior, A. N. 1960: ‘The Runabout Inference-Ticket’. *Analysis*, 21, pp. 38-39.
- 1960a: ‘Identifiable Individuals’. *Review of Metaphysics*, 13, pp. 684-696. Reprinted in Prior 1968.
- 1963-4: ‘Conjunction and Contonktion Revisited’. *Analysis*, 24, pp. 191-195.
- 1968: *Papers on Time and Tense*. Oxford: Clarendon Press.
- Prosser, Simon 2000: ‘A New Problem for the A-Theory of Time’. *The Philosophical Quarterly*, 50, pp. 494-498.
- Putnam, Hilary 1966: ‘The Analytic and the Synthetic’, in Feigl, H. and Maxwell, G. (eds.) *Minnesota Studies in the Philosophy of Science*, vol. III. Minneapolis: University of Minnesota Press.
- 1975: ‘The Meaning of “Meaning”’, in Gunderson, K. (ed.) *Language, Mind and Knowledge*. *Minnesota Studies in the Philosophy of Science*, 7. Minneapolis: University of Minnesota Press.

- 1983: ‘“Two Dogmas” Revisited’, in *Realism and Reason: Philosophical Papers*, vol. iii. Cambridge: Cambridge University Press.
- Quine, W. V. 1960: *Word and Object*. Cambridge, MA: MIT Press.
- 1961: ‘Two Dogmas of Empiricism’, in *From a Logical Point of View*. 2nd edition. Cambridge, MA: Harvard University Press.
- Read, Stephen 1994: *Thinking About Logic: An Introduction to the Philosophy of Logic*. Oxford: Oxford University Press.
- Recanati, François 1993: *Direct Reference: From Language to Thought*. Oxford: Blackwell.
- 1997: ‘The Dynamics of Situations’, in Dokic, Jérôme (ed.) *European Review of Philosophy*, vol. 2: *Cognitive Dynamics*. Stanford, CA: CSLI Publications. pp. 41-75.
- 2000: ‘Deferential Concepts: A Response to Woodfield’. *Mind and Language*, 15, pp. 452-464.
- Reichenbach, Hans 1956: *The Direction of Time*. Berkeley: University of California Press.
- Reynolds, Steven L. 1992: ‘Self-Recognition’. *The Philosophical Quarterly*, 42, pp. 182-190.
- Rozemond, Marleen 1993: ‘Evans on *De Re* Thought’. *Philosophia*, 22, pp. 275-298.
- Russell, Bertrand 1905: ‘On Denoting’. *Mind* 14, pp. 479-493.
- 1912: *The Problems of Philosophy*. Oxford: Oxford University Press.
- 1910-11: ‘Knowledge by Acquaintance and Knowledge by Description’. *Proceedings of the Aristotelian Society*, 11, pp. 108-128. Reprinted in *Mysticism and Logic*. London: George Allen and Unwin, 1917.
- 1919: ‘Descriptions’, in *Introduction to Mathematical Philosophy*. London: George Allen and Unwin Ltd. pp. 167-180. Reprinted in Martinich 1996.
- 1940: *An Enquiry into Meaning and Truth*. London: Allen and Unwin.
- 1957: ‘Mr. Strawson on Referring’. *Mind*, 66, pp. 385-389.
- 1985: *The Philosophy of Logical Atomism*. Edited and with an introduction by David Pears. La Salle, Illinois: Open Court.
- Salmon, Wesley C. 1984: *Scientific Explanation and the Causal Structure of the World*. Princeton N. J: Princeton University Press.
- Searle, John R. 1958: ‘Proper Names’. *Mind*, 67, pp. 166-173.
- 1984: ‘Intentionality and its Place in Nature’. *Dialectica*, 38, pp. 87-99.
- Shannon, Claude E. 1948: ‘A Mathematical Theory of Communication’. *Bell System Technical Journal*, 27, pp. 379-423, 623-56. Reprinted under the same title with an introductory essay by Warren Weaver. Champaign, Illinois: University of Illinois Press, 1949.
- Shoemaker, Sydney 1968: ‘Self-Reference and Self-Awareness’. *Journal of Philosophy*, 65, pp. 555-567.
- 1970: ‘Persons and Their Pasts’. *American Philosophical Quarterly*, 7, pp. 269-285.
- Sider, Theodore 2000: ‘Recent Work on Identity Over Time’. *Philosophical Books*, 41, pp. 81-89.
- Simons, Peter 2000: ‘Continuants and Occurrents’. *Aristotelian Society Supplementary Volume*, 74, pp. 59-75.
- Smith, E., and Medin, D. 1981: *Categories and Concepts*. Cambridge, MA: Harvard University Press.

- Smith, Quentin 1988: 'The Phenomenology of A-Time'. *Diálogos*, 52, pp. 143-153. Reprinted in Oaklander and Smith 1994.
- 1993: *Language and Time*. Oxford: Oxford University Press.
- Stalnaker, Robert C. 1984: *Inquiry*. Cambridge, MA/London: MIT Press/A Bradford Book.
- Strawson, P. F. 1950: 'On Referring'. *Mind*, 59, pp. 320-344.
- 1959: *Individuals: An Essay in Descriptive Metaphysics*. London: Methuen.
- Tarski, Alfred 1944: 'The Semantic Conception of Truth and the Foundations of Semantics'. *Philosophy and Phenomenological Research*, 4, pp. 341-375.
- Turetzky, Philip 1998: *Time*. London: Routledge.
- Turing, A. M. 1950: 'Computing Machinery and Intelligence'. *Mind*, 59, pp. 433-460.
- Villanueva, E., (ed.) 1991: *Philosophical Issues 1: Consciousness*. Atascadero, CA: Ridgeview.
- Vision, Gerald 2001: 'Flash! Fodor Splits the Atom'. *Analysis*, 61, pp. 5-10.
- Voltolini, Alberto 1997: 'Critical Notice of François Recanati's *Direct Reference*', in Dokic, Jérôme (ed.) *European Review of Philosophy*, vol. 2: *Cognitive Dynamics*. Stanford, CA: CSLI Publications. pp. 175-184.
- White, Morton G. 1976: 'The Analytic and the Synthetic: an Untenable Dualism', in Hook, S. (ed.) *John Dewey: Philosopher of Science and Freedom*. Westport, Connecticut: Greenwood.
- Wiggins, David 1987: *Needs, Values, Truth*. Oxford: Blackwell.
- Williams, Clifford 1992: 'The Phenomenology of B-Time'. *Southern Journal of Philosophy*, 30, pp. 123-137. Reprinted in Oaklander and Smith 1994.
- 1998: 'A Bergsonian Approach to A- and B-Time'. *Philosophy*, 73, pp. 379-393.
- Williamson, Timothy 1997: 'Sense, Validity and Context'. *Philosophy and Phenomenological Research*, 57, pp. 649-654.
- Wittgenstein, Ludwig 1958: *Philosophical Investigations*. Translated by G. E. M. Anscombe. 2nd edition. Oxford: Blackwell.
- 1958a: *The Blue and Brown Books*. Oxford: Blackwell.
- Woodfield, Andrew 2000: 'Reference and Deference'. *Mind and Language*, 15, pp. 433-451.
- Yourgrau, Palle (ed.) 1990: *Demonstratives*. Oxford: Oxford University Press.