



GARY KEMP

THIRD EDITION

WHAT IS THIS THING CALLED PHILOSOPHY OF LANGUAGE?

ROUTLEDGE

Praise for previous editions

- ‘This book is an outstanding pedagogical tool, which will be useful to anyone looking to gain a foothold in the subject. The second edition, which features new chapters on key figures, prominent topics, and recent developments in the field, is a substantial and welcome development of the excellent first edition.’

— **Brett Sherman**, University of South Carolina, USA

‘Will become the standard textbook for survey courses in the philosophy of language.’

— **Ernest Lepore**, Rutgers University, USA

‘To my mind this is the best introductory textbook for undergraduates looking to get a feel for the subject, without getting bogged down in advanced technical details. Gary Kemp covers all the traditional topics in the field and presents them in an accessible, engaging, and always rigorous style. Appended to each chapter are useful historical notes, a summary, a few questions, and some bibliographical recommendations for further research – a complete set of study aids that ought to be welcomed by students and teachers alike.’

— **Stefano Predelli**, University of Nottingham, UK

‘An easy, step by step journey through the classic themes of twentieth-century philosophy of language.’

— **François Recanati**, Institut Jean Nicod, France

‘Kemp has written a genuine introduction to the philosophy of language with beginning students in mind. Focusing on the issue of the meaning of natural language, he begins with a naive and, for students, very natural view of linguistic meaning. He then motivates and explains the distinctions, problems, solutions and development of the philosophy of language with the patience and understanding of a master teacher.’

— **Michael Losonsky**, Colorado State University, USA



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

What is this thing called Philosophy of Language?

- Philosophy of language explores some of the most abstract yet most fundamental questions in philosophy. The ideas of some of the subject's great founding figures, such as Gottlob Frege, Ludwig Wittgenstein and Bertrand Russell, as well as of more recent figures such as Saul Kripke and Hilary Putnam, are central to a great many philosophical debates to this day and are widely studied. In this clear and carefully structured introduction to the subject Gary Kemp explains the following key topics:
 - the basic nature of philosophy of language, its concepts and its historical development
 - Frege's theory of sense and reference; Russell's theory of definite descriptions
 - Wittgenstein's *Tractatus*, Ayer, and the Logical Positivists
 - recent perspectives including Kripke, Kaplan, Putnam, Chomsky, Quine and Davidson; arguments concerning translation, necessity, indexicals, rigid designation and natural kinds
 - the pragmatics of language, including speech-acts, presupposition and conversational implicature
 - puzzles surrounding the propositional attitudes (sentences which ascribe beliefs to people)
 - the challenges presented by the later Wittgenstein
 - contemporary directions, including contextualism, fictional objects and the phenomenon of slurs.

The third edition has been thoroughly revised throughout and includes a new chapter on Noam Chomsky's theory of Universal Grammar. In addition, the concluding chapter on modern directions in philosophy of language has been expanded to two chapters, which now cover crucial emergent areas of study such as slurs, conceptual engineering and experimental philosophy.

Chapter summaries, annotated further reading and a glossary make *What is this thing called Philosophy of Language?* an indispensable introduction to those teaching

philosophy of language and will be particularly useful for students coming to the subject for the first time.

Gary Kemp is Senior Lecturer in Philosophy at the University of Glasgow, UK. He has authored or edited various books and articles in the philosophy of language, including *Quine versus Davidson: Truth, Reference and Meaning* and *Quine's Philosophy: An Introduction*.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

What is this thing called?

The Routledge Philosophy *What is this thing called?* series of concise textbooks have been designed for use by students coming to a core area of the discipline for the first time. Each volume explores the relevant central questions with clear explanation of complex ideas and engaging contemporary examples. Features to aid study include text boxes, chapter summaries, study questions, further reading and glossaries.

Available:

What is this thing called Ethics? Second edition

Christopher Bennett

What is this thing called Metaphysics? Third edition

Brian Garrett

What is this thing called Philosophy of Religion?

Elizabeth Burns

What is this thing called The Meaning of Life?

Stewart Goetz and Joshua W. Seachris

What is this thing called Global Justice? Second edition

Kok-Chor Tan

What is this thing called Knowledge? Fifth edition

Duncan Pritchard

What is this thing called Metaethics? Second edition

Matthew Chrisman

What is this thing called Philosophy of Language? Third edition

Gary Kemp

For more information about this series, please visit: <https://www.routledge.com/What-is-this-thing-called/book-series/WITTC>

GARY KEMP

What is this thing called Philosophy of Language?

THIRD EDITION

 **Routledge**
Taylor & Francis Group
LONDON AND NEW YORK

● **Cover image:** © Getty Images

Third edition published 2024
by Routledge
4 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
605 Third Avenue, New York, NY 10158

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2024 Gary Kemp

The right of Gary Kemp to be identified as authors of this work has been asserted by them in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

First edition published by Routledge, 2013
Second edition published by Routledge, 2017

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

Names: Kemp, Gary, 1960 October 15- author.

Title: What is this thing called philosophy of language? / Gary Kemp.

Description: Third edition. | Abingdon, Oxon ; New York, NY :

Routledge, 2024. | Includes bibliographical references and index.

Identifiers: LCCN 2023037814 (print) | LCCN 2023037815 (ebook) |

ISBN 9781032426556 (hardcover) | ISBN 9781032426549 (paperback) |

ISBN 9781003363668 (e-book)

Subjects: LCSH: Language and languages—Philosophy.

Classification: LCC P107 .K458 2024 (print) | LCC P107 (ebook) |

DDC 121/.68—dc23/eng/20231201

LC record available at <https://lcn.loc.gov/2023037814>

LC ebook record available at <https://lcn.loc.gov/2023037815>

ISBN: 978-1-032-42655-6 (hbk)

ISBN: 978-1-032-42654-9 (pbk)

ISBN: 978-1-003-36366-8 (ebk)

DOI: 10.4324/9781003363668

Typeset in Berling
by codeMantra

CONTENTS

List of figures and tables	xiv
Preface to the third edition	xv

Introduction	1
• WHAT'S AHEAD	3
• EIGHT PREPARATORY NOTES	5
• COGNITIVE MEANING AND EXPRESSIVE MEANING	7
• MEANING AND FORCE	8
• CONTEXT-DEPENDENCE	9
• THE ROLES OF PROPOSITIONS	11
• COMPOSITIONALITY, STRUCTURE AND UNDERSTANDING	12
 1 Naïve semantics and the language of logic	 15
• NAÏVE THEORY: SINGULAR TERMS, PREDICATES AND REFERENCE	15
• TRUTH AND MEANING FOR ATOMIC SENTENCES	21
• LOGICAL SYNTAX AND LOGICAL OPERATORS	23
• GENERALISING AT THE LEVEL OF SINGULAR TERMS AND PREDICATES	26
HISTORICAL NOTES	26
CHAPTER SUMMARY	27
STUDY QUESTIONS	27
PRIMARY READING	28
NOTES	28
 2 Fregean semantics	 29
• TWO PROBLEMS FOR NAÏVE SEMANTICS	29
• THE SENSE-REFERENCE DISTINCTION	32
• THE DISTINCTION EXTENDED	33
• COMPOSITIONALITY AGAIN; THE REFERENCE OF A SENTENCE	34
• APPLYING THE THEORY	36
• SUBSTITUTIVITY AND EXTENSIONALITY	38
• THE ANALYSIS OF PROPOSITIONAL ATTITUDES	39
• THE OBJECTIVITY OF SENSE	41
• PREDICATE REFERENCE AND THE CONCEPT HORSE PROBLEM	43
• FURTHER DISCUSSION: THE CONTEXT PRINCIPLE	45
HISTORICAL NOTES	46

CHAPTER SUMMARY	47
STUDY QUESTIONS	49
PRIMARY READING	50
SECONDARY READING	50
NOTES	51
3 Russellian semantics	52
• THE TASK FOR RUSSELL	52
• THE THEORY OF DEFINITE DESCRIPTIONS	54
• APPLYING THE THEORY OF DESCRIPTIONS	57
• NAMES AS DISGUISED DEFINITE DESCRIPTIONS	60
• KNOWLEDGE BY ACQUAINTANCE AND KNOWLEDGE BY DESCRIPTION	61
HISTORICAL NOTES	65
CHAPTER SUMMARY	65
STUDY QUESTIONS	66
PRIMARY READING	67
SECONDARY READING	67
NOTES	67
4 Russell's theory of judgement, the early Wittgenstein and logical positivism	68
• PROPOSITIONS, FACTS AND RUSSELL'S THEORY OF JUDGEMENT	69
• THE <i>TRACTATUS LOGICO-PHILOSOPHICUS</i>	72
• VERIFICATIONISM I: AYER	77
• VERIFICATIONISM II: CARNAP'S LOGICAL EMPIRICISM	82
• THE VIENNA CIRCLE AND THE PROTOCOL DEBATE	85
HISTORICAL NOTES	88
CHAPTER SUMMARY	89
STUDY QUESTIONS	91
PRIMARY READING	92
SECONDARY READING	92
NOTES	92
5 The late Wittgenstein	93
• LANGUAGE GAMES	93
• FAMILY RESEMBLANCE, TOOLS AND CITIES	96
• TO FOLLOW A RULE I	98
• TO FOLLOW A RULE II	99
• PRIVATE LANGUAGE	101
HISTORICAL NOTES	102
CHAPTER SUMMARY	103
STUDY QUESTIONS	103
PRIMARY READING	104
SECONDARY READING	104
6 Quine's philosophy of language	105
• QUINE'S NATURALISM	105
• THE FIELD LINGUIST	107
• INDETERMINACY	109

•	MEANING AND ANALYTIC TRUTH	111
•	THE ARGUMENT OF 'TWO DOGMAS OF EMPIRICISM'	112
•	REGIMENTATION; ANALYSIS VERSUS REPLACEMENT	114
•	FURTHER DISCUSSION: ONTOLOGICAL RELATIVITY, AND INSCRUTABILITY AND INDETERMINACY DISTINGUISHED	116
•	THE PLACE OF NATURALISM	118
	HISTORICAL NOTES	119
	CHAPTER SUMMARY	119
	STUDY QUESTIONS	120
	PRIMARY READING	121
	SECONDARY READING	121
	NOTE	121
7	Kripke on naming and necessity	122
•	NECESSITY, POSSIBILITY AND POSSIBLE WORLDS: A PRIMER	122
•	THE DESCRIPTIVIST PARADIGM	124
•	KRIPKE'S OBJECTIONS TO THE DESCRIPTION THEORY OF PROPER NAMES	125
•	RIGID DESIGNATION	127
•	FIXING THE REFERENCE I: CAUSAL CHAINS	129
•	FIXING THE REFERENCE II: DESCRIPTIONS	130
•	LINGERING ISSUES FROM RUSSELL AND FREGE	132
•	FURTHER DISCUSSION: INTENSIONAL SEMANTICS	132
	HISTORICAL NOTES	134
	CHAPTER SUMMARY	135
	STUDY QUESTIONS	136
	PRIMARY READING	137
	SECONDARY READING	137
	NOTES	137
8	Context-dependence, indexicality and natural kinds	139
•	INDEXICALS AND DEMONSTRATIVES	139
•	PUTNAM ON NATURAL KIND TERMS AND ESSENCE	141
•	IS MEANING IN THE HEAD?	143
•	THE ACTUAL WORLD AS A CONTEXT	144
•	TWO-DIMENSIONALISM: CONTEXT OF UTTERANCE VERSUS CIRCUMSTANCE OF EVALUATION	146
•	FURTHER DISCUSSION: RIGID DESIGNATION AGAIN	148
•	THE INDISPENSABILITY OF INDEXICALS	150
•	INDEXICALS AND FREGEAN SENSE	151
	HISTORICAL NOTES	152
	CHAPTER SUMMARY	153
	STUDY QUESTIONS	154
	PRIMARY READING	155
	SECONDARY READING	155
	NOTES	156
9	Pragmatics	157
•	MOOD AND FORCE REVISITED	157
•	SPEECH-ACT THEORY	158
•	IMPLICATURE	162

• SOME APPLICATIONS OF THE CONCEPT OF IMPLICATURE	164
• PRESUPPOSITION; STRAWSON'S AND DONNELLAN'S OBJECTIONS TO RUSSELL'S THEORY OF DESCRIPTIONS	166
• METAPHOR	168
HISTORICAL NOTES	170
CHAPTER SUMMARY	171
STUDY QUESTIONS	173
PRIMARY READING	173
SECONDARY READING	173
NOTES	173
10 Davidson's philosophy of language	175
• METHODOLOGY	176
• THE GENERAL FORM OF A THEORY OF MEANING	176
• DAVIDSON'S WAY: T-SENTENCES	178
• AN EXAMPLE OF A THEORY OF MEANING	180
• FURTHER DISCUSSION: QUANTIFIERS AND VARIABLES	182
• THE EMPIRICAL CONFIRMATION OF A THEORY OF MEANING: RADICAL INTERPRETATION	183
• THE PRINCIPLE OF CHARITY AND THE INTERDEPENDENCE OF BELIEF AND MEANING	185
• FURTHER DISCUSSION: EVENT SEMANTICS AND INDIRECT SPEECH	187
HISTORICAL NOTES	189
CHAPTER SUMMARY	190
STUDY QUESTIONS	191
PRIMARY READING	192
SECONDARY READING	192
NOTES	192
11 The propositional attitudes	193
• EXTENSIONALITY REVISITED	193
• REFERENTIAL OPACITY AND FREGE ON THE ATTITUDES	194
• FURTHER DISCUSSION: MULTIPLE HYPER-INTENSIONAL EMBEDDING	197
• <i>DE RE</i> AND <i>DE DICTO</i> NECESSITY	198
• <i>DE RE</i> AND <i>DE DICTO</i> BELIEF	199
• RALPH'S PREDICAMENT	200
• BELIEF ATTRIBUTIONS AND EXPLICIT INDEXICALS; BELIEF <i>DE SE</i>	203
• AN IMPLICIT INDEXICAL ELEMENT	204
• DIRECT REFERENCE, THE ATTITUDES AND THE SEMANTIC <i>DE RE</i>	206
HISTORICAL NOTES	208
CHAPTER SUMMARY	209
STUDY QUESTIONS	210
PRIMARY READING	211
SECONDARY READING	211
NOTES	211

12 Chomsky's science of language and Universal Grammar	212
• PLATO'S PROBLEM AND BIOLOGY	214
• INTERNAL LANGUAGE AND EXTERNAL LANGUAGE	216
• ELEMENTARY PHRASE STRUCTURE GRAMMAR	217
• TRANSFORMATIONAL GRAMMAR; TRACE THEORY	219
• THE TOTAL LANGUAGE FACULTY	221
• KNOWLEDGE, RULES, PRINCIPLES AND PARAMETERS	223
• MEANING, REFERENCE AND CONCEPTS	224
• LANGUAGE, COMMUNICATION, BEHAVIOURISM; PERFORMANCE VS COMPETENCE	226
• A RECENT DEVELOPMENT, AND THE POSSIBLE ORIGINS OF LANGUAGE	228
HISTORICAL NOTES	229
CHAPTER SUMMARY	230
STUDY QUESTIONS	232
PRIMARY READING	233
SECONDARY READING	233
NOTE	233
13 Modern directions I	234
• ASSERTION	234
• CONTEXT-RELATIVITY	237
• FICTIONAL OBJECTS	242
• INFERENTIALISM	245
CHAPTER SUMMARY	246
STUDY QUESTIONS	248
PRIMARY READING	248
SECONDARY READING	249
NOTE	250
14 Modern directions II	251
• SLURS AND PEJORATIVES	251
• CONCEPTUAL ENGINEERING	259
• X-PHI (EXPERIMENTAL PHILOSOPHY)	268
CHAPTER SUMMARY	271
STUDY QUESTIONS	273
FURTHER READING (PRIMARY AND SECONDARY)	274
NOTES	275
Glossary	277
Bibliography	299
Index	307

• figures and tables

FIGURES

1.1	The proposition expressed by 'Mars orbits the sun'	22
2.1	Sense and reference	32
2.2	The relation of sense to reference	33
2.3	Fregean semantics for one-place atomic sentences	35
4.1	Acquaintance	71
4.2	Judgement	72
10.1	The two factors responsible for holding a sentence true	186
12.1	'Some person took the last one'	218
12.2	'Sally is from San Francisco'	218
12.3	'The tall girl won the race' (I)	219
12.4	'The tall girl won the race' (II)	219
12.5	The Basic Structure of Universal Grammar (UG)	221

TABLES

4.1	Truth-table for 'not', 'or' and 'and'	75
8.1	Two-dimensional diagram for 'water= H ₂ O'	148
9.1	Mood and force	158
9.2	Force distinctions within a speech act	160
11.1	The sense and reference of a term as used in three contexts	198
13.1	Truth-table for 'and'	245
13.2	Inference-rules for 'and'	245

• preface to the third edition

Many students approaching the philosophy of language for the first time are baffled by it. They are baffled in a way that they are not baffled by moral philosophy, political philosophy or the theory of knowledge. Partly because some of the hardest steps come right at the beginning, they find it excruciatingly difficult to find their feet with the subject, as if they were learning to surf or to ski.

The situation is not helped by the fact that most of the primary materials are written at the highest philosophical level, with many technical terms and principles assumed to be already understood that will be Greek to the neophyte. This textbook is meant to help the beginning student into the subject. There are some excellent textbooks already but few for the genuine beginner. I'll name one that is pitched at roughly the same level as this: William Lycan's *Philosophy of Language: A Contemporary Introduction* (Routledge, third edition, 2018). The book is outstanding, but there are large differences between his book and this one. His book is arranged topically; it is full of names and isms, and the problems, replies and counter-replies come thick and fast. By contrast, this book is for the most part theory-based and often theorist-based, more concerned to keep one's eyes fixed on the larger and deeper issues and themes, with fewer names mentioned, a smaller range of problems considered and – crucially – a slower pace. There is no getting around the difficulty of the first steps, but this book attempts to immerse the student into a few paramount theories and their authors – Frege, Russell etc. – in order to get them used to thinking within the author's point of view, and to get them to see why one would think as the author does.

In addition, and perhaps most importantly, I hope the student will come away with a coherent picture of the history of the subject. The history will not be thorough. The sketch is only of the main lines, and only of those lines – if such there be – which are for the newcomer both interesting and relatively easy to learn. There is nothing or precious little of Burge, Church, Dummett, Evans, Fodor, Geach, D. Lewis, Millikan, Montague, Schiffer, Sellars, Stalnaker and Tarski, and only selected bits of Carnap, Ayer, Grice, Searle and Kaplan. Nothing or precious little of teleosemantics, truth-maker semantics, conceptual role semantics, situation semantics, game-theoretic semantics, dynamic semantics, intention-based semantics, semantical

minimalism, expressivist semantics, realism versus anti-realism, relevance theory or the theory of truth. I hope that students will be able to pursue those topics once they've got their feet wet in the philosophy of language by studying this book and the authors it discusses.

Likely criticisms of the book are that it is too much weighted towards the drier end of the subject, and towards the history of the subject. These are connected. To take the first one first, its connections with the philosophy of mathematics and of mind, with metaphysics and epistemology, have historically driven the subject – indeed, one might say, have constituted its main reason for existing. Such earlier figures as Frege, Russell, Wittgenstein and Carnap, and later figures such as Quine, Davidson and Kripke, thought of philosophical reflection on language as being the key towards progress on those fundamental subjects. And there is no getting around the fact that the most interesting ideas in this sector are abstract and difficult. Some students will gravitate away from this end of the subject, towards the more political and ethical end that has grown in the past thirty years, but it is eminently arguable that in order to speak with penetration at that end, one needs a foundation in the drier end, to use with finesse such concepts as reference, truth-condition, speech-act and so on. The complaint about history, meanwhile, boils down to a philosophical difference about how to teach the philosophy of language. In my view the subject is not like geometry or mathematical logic, which normally are taught in complete detachment from their histories. The subject of philosophy of language is too contentious, as evinced by the broad range of isms and theories mentioned in the last paragraph. I think that one is better placed to review more recent material and controversies – a taste of which is included at the end of this book – if one first has a solid understanding of the basics of Frege, Russell and others.

Another likely criticism is that, outside the study questions, it is not very critical. I admit but do not repent. I feel that a big mistake that is often made in teaching the philosophy of language is to criticise a position almost the moment it's on the table. In my experience, it takes some time for a position to sink in, especially to see why one might hold it. If the teacher criticises the view from the get-go, the student is unlikely to think it worth spending time on it – aside from exam-taking purposes – and may wonder why the position is being taught if it is so obviously full of holes. At the risk of betraying my Californian roots, I want to put a more positive spin on the material.

Thus, after a modicum of stage-setting in the form of an ahistorical introduction and first chapter, this book considers Frege, Russell, the early Wittgenstein and the Logical Empiricists, including Ayer's early view and parts of Carnap; then the later Wittgenstein, Quine, Kripke, elementary possible worlds semantics, Putnam, indexicals and the basics of pragmatics – Austin, Grice and Searle – then Davidson, then the propositional attitudes centring on the problems raised by Frege and Quine but with a modern dimension as well, then Chomsky's Universal Grammar, with an eye on the more philosophical aspects. In the last two chapters, I review some of the principal developments in recent years including some subjects at the more political and ethical end of the subject: assertion, context-relativity, fictional objects,

inferentialism, slurs/pejoratives, conceptual engineering, and X-Phi (Experimental Philosophy).

The chapters are supposed to be read in sequence, but the chapters are written so as to be relatively self-contained; it is not impossible to depart from the order in which the chapters are presented. Two chapters – Chapter 8 on possible worlds and indexicals and Chapter 11 on the propositional attitudes – are slightly more difficult than the others (and Chapter 4 is more historical than the others). You won't hurt my feelings if you skip them (or any others for whatever reason). Since not everyone will read every chapter, there is some repetition of points, conceptual and historical.

At the end of each chapter are four items (with the exceptions of Chapters 13 and 14, which lack the first item on the list):

- 1 Some historical notes, including some gossip material.
- 2 A chapter summary.
- 3 Study questions, which are not just questions for which the answers are present in the chapter but are designed to get one to think more reflectively and critically about and with the material.
- 4 Suggested further reading; I take it every reader knows how to obtain the relevant entries in the *Stanford Encyclopedia of Philosophy* and the *Internet Encyclopedia of Philosophy*, so I save a little space by not listing them except when I think what they offer is especially valuable. I list the crucial references from, for instance, Frege, as 'Primary reading'; and in most chapters I've also listed some items under 'Secondary reading'.

At the end, you'll find a glossary of terms introduced in the book.

I stress that the primary reading is essential to any serious course using this book. I don't think any textbook can *replace* the original works; as mentioned above, the subject is too contentious to be like chemistry or calculus, with only historians interested in the original texts. This book, I hope, will serve as an initial exposure to the writers, issues and arguments in the philosophy of language, and as a concise map, serving to orient the reader through the primary reading.

In the third edition: Chapter 12 (Chomsky's Universal Grammar) is new; plus there is a substantial expansion of what was Chapter 12 ('Modern directions') into two, Chapters 13 and 14. I have revamped the order of chapters from Chapter 5 on, into a more straightforward chronological order, and added sections to the chapter on Davidson. And there are smaller-scale adjustments, many at the urging of reviewers for Routledge; I thank them. Also I single out David Lumsden and Michael Lumsden for thanks, and also the many students who have commented.

Dr Gary Kemp
September 2023



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

• introduction

Western philosophy has been explicitly concerned with language since the early twentieth century and implicitly for much longer. Indeed, for much of the twentieth century – the period of the so-called ‘linguistic turn’, discussed explicitly in Chapter 4 – some people were convinced that the philosophy of language just *was* philosophy. Why? What exactly has language got to do with philosophy, or philosophy with language?

Here are some preliminary ideas, ones that have tended to motivate the philosophy of language:

- As philosophers we ask ‘What is justice?’ or ‘What is the nature of justice?’; but we can also ask ‘What is the meaning of the word “justice”?’ Questions about the essences, or natures of things, can seemingly be transformed into questions about semantics or the meaning of words. Some see such shifts as philosophical progress (others reckon that it trivialises philosophy).
- Language expresses thought; it mirrors thought. The study of language is one way to study thought – its character, its structure and its relation to the world. ‘The limits of my world are the limits of my language’, once said Ludwig Wittgenstein. Furthermore, language being a social thing, the study of language is the study of the ways in which our thoughts are shaped and moulded by society. And unlike thought itself, language is out in the open, open to objective scrutiny in ways that thought is not.
- Language represents the world. It mirrors it. Thus the study of the more general or abstract features of language might be thought to reveal the more general or abstract features of the world.
- The study of language is itself partly a philosophical enterprise: language exists in the real world and is thus open to scientific scrutiny, but it is not immediately obvious what a scientific theory of language would be like or what exactly the relevant data would be. So we have to reflect in an *a priori* way before we know what sorts of questions to ask.
- The analysis of language – especially the theory of meaning as informed by logic – enables us to understand what *clarity* is. Since one of the defining features of philosophy is its struggle to clarify difficult or otherwise problematic ideas, the enterprise assists philosophy in its task.

Each of these might be denied or quibbled with in various ways. However, they are collectively plausible enough to motivate the philosophical investigation of language.

From its beginnings in the latter half of the nineteenth century, the theoretical study of language has standardly been divided into three main areas: **syntax**, **semantics** and **pragmatics**.

Syntax deals with relations between symbols and symbols, between signs and signs.

It is most fundamentally concerned with **grammar**: among other things, it aims to discover those basic principles that determine whether a given string of signs is grammatically 'well formed'.

Semantics deals with relations between symbols and what they mean, express or are about. How to characterise its fundamental concern is more contested, but a dominant tradition conceives it as formulating a system of rules which determine the **truth-condition** of an arbitrary declarative sentence.

Pragmatics is yet more various in its aims, but, broadly speaking, it is concerned with the **use** of sentences – given that a certain sentence has certain basic semantic properties, what sorts of act can be accomplished in actual communicative situations by uttering it?

These divisions should by no means be regarded as absolute. Considerations from one sub-discipline may have repercussions for another, disputes break out over which of two sub-disciplines is the more fundamental or what it means to call it 'fundamental', and sometimes the rationale is called into question for maintaining the tripartite division in the first place.

As intimated by the fourth bullet point in the list above, there is a scientific discipline known as linguistics. What is the relation between that science and the philosophy of language?

There is no single, neat way of distinguishing them. The distinction is partly one of degree rather than of kind, and partly an historical matter of institutional arrangements – those which explain, for example, the fact you can study formal logic in computer science, mathematics or philosophy. But we can say some further things.

First, to a much greater degree than linguistics, philosophy is concerned with the interface of language with problems of knowledge, metaphysics, ethics, politics and aesthetics.

Second, linguistics is much more concerned with empirical facts, with real, contingently existing languages in all their diversity, with the perplexing puzzles thrown up by actual languages in the wild. Philosophy tends to be concerned with the *a priori* side of things, perhaps risking jokes about armchairs and port.

And, lastly, we can divide linguistics into the theoretical side and the applied side. Philosophy of language has more in common with the theoretical side, especially in its interest in the theory of **meaning**. The theory of meaning, broadly speaking, is the same as semantics, both in linguistics and as pursued by philosophers. But in philosophy the domain tends to be narrower on the one hand and, we like to think,

deeper on the other, than that of theoretical linguistics. Standard philosophical and linguistic theories employ the concepts of **reference** – of *aboutness*, or *of-ness* – of truth, **cognitive content** and many others. But questions of the status of those concepts, of their justification, place and nature, are philosophical. An imprecise but useful analogy is with the relation of mathematics to the philosophy of mathematics: in mathematics you use things such as numbers, fractions and functions, methods such as algebra, trigonometry and the calculus. But when you ask ‘What is the ultimate basis of algebra, trigonometry and the calculus? How do we know about them? What *are* numbers, fractions and functions? Do they really exist along with stones and cats?’ you are engaged in the philosophy of mathematics. Philosophy of language is likewise more reflective than linguistics, more likely to adopt a ‘meta-perspective’ on theories of meaning (sometimes philosophers say that what they are doing is ‘meta-semantics’).

Until much later, when we introduce the views of Noam Chomsky, we will say comparatively little about syntax as linguists understand the term, covering only the rudiments needed to convey the various approaches to the theory of meaning. We will have more to say about pragmatics in the middle of the book, not only for its intrinsic philosophical interest and importance, but also because pragmatic considerations tend to relieve some of the lingering worries one might have about the main ideas of semantics or the theory of meaning.

• WHAT’S AHEAD

This book begins by exploring certain classical theories of meaning – classical in the sense that they continue to serve as the reference points for more recent theories of meaning, and as the kernel of philosophy of language with which most philosophers in other fields are familiar. Although it has its precursors in such figures as Plato, Antoine Arnauld and Pierre Nicole (authors of the *Port Royal Logic*), and John Locke, the philosophy of language reached its first maturity relatively recently, with the work of Gottlob Frege (1848–1925) and Bertrand Russell (1872–1970), both of whose ideas are very much part of the contemporary scene; these are the subjects of Chapters 2 and 3. Those chapters are prefaced by Chapter 1, a discussion of a relatively simple theory that appeals to common sense; we call it ‘Naïve Semantics’. In describing the theory, we take the opportunity to introduce some elementary logical notions – singular term, predicate, truth-functional connective and so on (those who have already studied logic can easily skip over some of the chapter).

Chapter 4 concentrates on a famous movement – ‘Logical Positivism’ or ‘Logical Empiricism’ – that happened in the 1920s and 30s and dominated philosophy of language and analytic philosophy generally until the 1950s. We will discuss two of its leading figures, Rudolf Carnap (1891–1970) and A. J. Ayer (1910–89), plus the early work of Ludwig Wittgenstein (1889–1951); his *Tractatus Logico-Philosophicus* (1921) was a powerful influence. In Chapter 5 we encounter a famous note of scepticism concerning the theoretical ambitions of the philosophy language that arose in the 1950s, namely Ludwig Wittgenstein’s later philosophy as expressed

in his *Philosophical Investigations* (1953). Chapter 6 deals with another angle of doubt for the prospects for a theory of meaning, but with a positive theoretical programme for addressing many of its main concerns: the programme of W. V. Quine (1908–2000).

In Chapter 7 we get back on the bus of orthodoxy, discussing a modern alternative that arose in the 1960s and 70s, largely in response to Frege and Russell but also in opposition to Logical Positivism, namely the direct reference theory popularly associated with Saul Kripke (1940–2022); the theory has various ramifications for other branches of philosophy, especially for metaphysics and epistemology. Chapter 8 deals with a topic that grew in stature from the late 1960s on, namely the context-variability of what are known as **indexicals**, such as ‘I’ and ‘here’; we’ll consider the topic through the lens provided by David Kaplan, and its philosophical ramifications as articulated by Hilary Putnam (1926–2016) and Frank Jackson.

Pragmatics, the use of language, is the subject of Chapter 9, especially as first presented by J. L. Austin (1911–60), H. P. Grice (1913–88), and John Searle in the 1960s and 70s. Chapter 10 discusses a celebrated theory that arose in the 1970s and 80s due to Donald Davidson (1917–2003), that attempts to describe the facts underlying meaning and reference without simply helping itself to those concepts. Chapter 11 delves more thoroughly into a puzzling area that inspired many of the main moves in both Frege’s and Russell’s philosophies of language, and which continues both to puzzle and inspire students at all levels. This is the semantics of **propositional attitudes**, such as ‘Darwin believed that human beings and gorillas have a common ancestor’.

Chapter 12 presents the views of Noam Chomsky largely as a contribution to philosophy, rather as than a contribution to the science of linguistics. A sketch of ‘**Universal Grammar**’ will be provided, but only as a necessary background for Chomsky’s most striking philosophical pronouncements, such as that rationalism rather than empiricism is correct, that all human languages are fundamentally the same, and that it’s not the case that the function of language is communication.

Our final two chapters explore topics which are currently receiving a lot of attention. Chapter 13 covers **assertion**, **fictional objects**, **context-relativity** (the idea that meaning shifts with context even for ordinary terms such as ‘know’) and **inferentialism** (the idea that inference rather than reference determines meaning); Chapter 14 covers **slurs**/pejorative language, demeaning language which has a certain political or ethical dimension (especially the language of gender and race but also more generally); **X-phi** or experimental philosophy (surveys etc. rather than the philosopher’s intuitions as philosophical data); and **conceptual engineering**. Other subjects that might have been considered include lying, propaganda, silencing, dynamic semantics, hate speech and free speech; but there are space constraints.

By the end of the book, you should know the most celebrated names and theories in the philosophy of language and be conversant in its paradigmatic arguments, theories and kinds of criticism. But if you really want to know your stuff, it is essential that you study the works listed as primary reading. Of course, there is *much* more to the field; look to the historical notes and initial secondary reading for guidance.

We begin with a few foundational points, and, crucially, we begin to introduce some special terminology. Some of what follows is eminently debatable, but it will help to have certain structures before us that, in one sense or another, are more or less accepted by all of the main figures that we discuss except perhaps Wittgenstein and Quine.

● EIGHT PREPARATORY NOTES

- 1 You may have noticed already that key technical or theoretical terms, when first used, are printed in **boldface**; these terms are in the glossary.
- 2 A few later sections are marked 'Further discussion'; naturally, they are philosophically interesting, but they are harder than our main concerns or just less essential to them.
- 3 Ask yourself: how many words does the following sentence contain?

Your dog bit my dog.

Trick question! It depends on what we mean by a 'word'. The sentence has two **tokens** of the **type** 'dog'. So the sentence has five words counted as tokens, four counted as types. We also say that the word-type 'dog' *occurs* twice in the sentence.

- 4 A word about what 'is' is. An argument-parody concerning the great blues musician Ray Charles runs: 'God is love; and love is blind; but Ray Charles is blind; therefore Ray Charles is God'. Maybe Ray Charles is indeed God, but the reasoning doesn't support that conclusion. The fallacy is to interpret 'is', at every occurrence, as indicating *identity* – as in '=' or 'is the very same thing as' or 'is identical with' – rather than as indicating **predication**, as in 'The cat is hungry.' Charles is blind, but he is not *identical* with blindness. Stevie Wonder is also blind, but he is not identical with blindness either. Indeed, if we thought otherwise, then according to the symmetry of identity (that if $a=b$ then $b=a$) and its transitivity (that if $a=b$ and $b=c$, then $a=c$), Stevie Wonder and Ray Charles would be the same man! But they are two men, not one. So we must distinguish the two senses. In this book, we will tend to use the equals sign, the identity sign '=', for the 'is' of identity, reserving 'is' for predication.
- 5 Philosophers of language use terms of ordinary language, but we try to mean them in a more specific way than the sometimes slipshod way they are used ordinarily. Especially important is the way the ordinary term 'reference' is employed in the philosophy of language. Consider the relation between 'Boston' and Boston. Without violating ordinary usage, we can say that that word *designates, labels, means, denotes, indicates, picks out, mentions, names, is the name of, is about, stands for, has the content of, signifies* the city. Maybe these differ slightly, but what is arguably common to them is the relation that we call 'reference'. We say that the word *refers* to the city: Boston is the *referent* of 'Boston'.

- 6 We need to be clear about the use of *quotation* marks to talk *about* language. For example, the following are true:

- (a) Boston is a city on the east coast of the USA.
- (b) 'Boston' contains six letters.

(a) says something about a city, whereas (b) says something about a *word*, the name of a city. As we say, (a) **uses** the word 'Boston', whereas (b) only **mentions** the word. Strictly speaking, (b) does not say anything about Boston, only about its name. The following sentence both uses and mentions 'Boston':

- (c) 'Boston' refers to Boston.

We can go further, nesting quotation marks within quotation marks. For example:

- (d) "'Boston'" refers to 'Boston'.

Compare (d) with (c). Whereas (c) says that a name of Boston refers to Boston, (d) says that a name of a name of Boston refers to a name of Boston.

In some contexts, this small detail can make all the difference. If we are going to talk about language, then we had better make sure we know which bit of language we are talking about, and we had better make sure we are not talking about the world when we mean to talk about language, or language when we mean to talk about the world. Serious philosophical errors have been made precisely by being sloppy over this (the logician-philosophers Quine and Kurt Gödel (1906–78) famously took none other than Russell to task over it).

A word to the wise: when writing philosophy essays, certainly essays in the philosophy of language, be careful about using quotation marks for any other purpose. That is, try to avoid the use of 'scare quotes', as I just used them, in speaking of scare quotes. For I meant to say something about scare quotes, not about 'scare quotes'. At its worst, the use of scare quotes is an evasive way of using a bit of language while at the same time distancing yourself from it, leaving your reader wondering whether you quite stand by what you say.

- 7 We will often speak of the **truth-conditions** of sentences or statements, and of their truth-value. By the former we mean 'the circumstances under which the statement is true', or 'the set of circumstances under which the statement is true'. Thus the truth-condition of 'Spot is hungry' is simply that Spot is hungry. Its truth-value is truth if Spot is indeed hungry, and falsity if Spot is not hungry. The truth-value is not the *value of the statement's being true*, or *how much truth it has in it*. It is not rocket science, but it does make the language we employ for doing the philosophy of language more precise and stable, better for making sure we understand one another.
- 8 The notion of an analytic sentence is appealed to and used by many philosophers in this book. The basic idea is that an analytic truth is one that can be known to be true *just from the meaning of the sentence* (similarly, an analytic falsehood is one can that be known in that way to be false). One can't understand an analytic truth without accepting it. To take the hackneyed example:

(1) No bachelor is married.

Anyone familiar with those words, who knows what they mean (and the significance of assembling them in this way), can see on that basis that the sentence is true; one can see that it would be *contradictory* to deny it; a married bachelor is impossible, a contradiction in terms. One does not have to take a poll of bachelors to know that they are not married. For *non-analytic* truths, *synthetic* truths, the case is otherwise:

(2) King Charles is not a bachelor.

Or:

(3) No bachelor is tidy.

To establish these as true (if they are true) it is not enough merely to reflect on the meaning of the words involved; one needs empirical information, hard data of some sort concerning King Charles and bachelors.

There is a connection via the notion of *synonymy* with the notion of logical truth. ‘Bachelor’ and ‘unmarried man’, we assume, are synonyms, mean the same. One can replace any occurrence of the one with an occurrence of the other (except where it is mentioned rather than used) without disturbing the meaning of the overall sentence. So from (1), replacing ‘bachelor’ with ‘unmarried man’, we derive:

(4) No unmarried man is married.

This is also an analytic truth, but it is also explicitly a truth of logic. If a truth can be transformed into logical truth by substituting synonyms, then it is an analytic truth.

● COGNITIVE MEANING AND EXPRESSIVE MEANING

Consider the following pair of sentences:

Karen’s small cat died.

Karen’s wee pussycat passed away.

Do these sentences mean the same or not? In one sense, they do; in another sense, they don’t. What they have in common is generally called **cognitive meaning**. The two sentences can be used to *convey the same objective fact* or the *same information*, namely the death of a certain feline. They have the same truth-condition: in any conceivable circumstance, they are either both true or neither is (assuming that ‘Karen’ refers to the same person).

The other dimension is the domain of rhetoric and spin: the same information, it seems, can be conveyed in different ways, conveying different subjective attitudes or feelings about it. We call this **expressive meaning**. Not all language is equally

possessed of expressive meaning; in our example, the first sentence is relatively flat or colourless in comparison with the second. The expressive meaning of scientific language, especially when mathematical, seems to be minimal or absent.

Until Chapter 14, the lion's share of our concern will be with cognitive meaning, as this is what links up most directly with enduring philosophical issues of epistemology and metaphysics. Except where ambiguity threatens, we will speak simply of meaning, as short for 'cognitive meaning'.

● MEANING AND FORCE

The most conspicuous purpose of language, if not its only purpose, is *communication*. Communication is normally achieved by means of *linguistic acts*, or **speech-acts**.

With some exceptions, one performs a speech act, says something, by uttering a complete sentence, or by uttering something that is intended in such a way as to be equivalent, for the purposes at hand, to a complete sentence. For example, if questioned 'Are you a student?', you might answer 'Yes'. What you say is not a complete sentence, but it is equivalent for the purposes at hand to 'I am a student'. (The obvious exceptions are greetings such as 'Hello!' and exclamations and the like, such as 'Crikey!')

Consider now the following sentences (pretend they are addressed to yourself):

- (5) You are going to eat raw fish.
- (6) Are you going to eat raw fish?
- (7) Eat raw fish!

The first is a sentence in the **declarative** (also called indicative) **mood**, the second in the *interrogative mood*, the third in the *imperative mood*. They have a certain something in common, namely the idea *that you are going to eat raw fish*. The first would normally be used to assert, or say that, you are going to eat raw fish; the second would normally be used to ask whether you are going to eat raw fish; the third would normally be used to suggest, command or enjoin you (the listener) to eat raw fish.

What is this thing that these have in common, which we can express by means of the clause *that you are going to eat raw fish*. We will say that this common element is a **proposition** – the proposition that you are going to eat raw fish. As used in a normal context, (5) to (7) all express this proposition, but the forms are normally used to (5) assert that it is true; (6) ask whether it is true; (7) suggest or enjoin that it be made true. We sum this up by saying: (5) to (7) are used to attach a different **force** to the self-same proposition. (5) is normally used to attach **assertoric** force to it, (6) *interrogative* force to it and (7) *imperative* force to it. Of course, one can *utter* 'You are going to eat raw fish' without actually asserting anything, as a stage actor might do. By varying one's intonation, one could ask a question using that form of

words – ‘You are going to eat raw fish?’, a friend might ask incredulously. And so on. Whether one actually attaches a given type of force to a proposition is determined, in typical circumstances, by the intention with which one speaks and the context in which one speaks, not just by the form of words uttered. Still, each of the three *grammatical moods* exemplified by (5) to (7) is normally used to express a characteristic force: it is by using the appropriate mood that we typically make it known which force we attach to the proposition expressed. Mood is a feature of grammar or, more technically, of syntax; force is a feature of pragmatics.

From now on, we are mostly going to *ignore* non-declarative sentences. Thus when speaking of ‘sentences’ we are usually speaking of declarative sentences. We need to say more about the relation of the proposition to the sentence. Consider:

Snow is white.

La neige est blanche.

Schnee ist weiss.

These declarative sentences, we should naturally say, are *synonymous*, in the sense that they are correct translations of each other. They mean the same thing, namely that snow is white. They all *express* the proposition that snow is white: the meaning of a sentence is the proposition it expresses. The common element, the proposition, is *not* any particular form of words – not a sentence, and not a clause of English such as ‘that snow is white’. Another way to put the same point is that these sentences all have the same *content*.

Notice that we are speaking of a *meaning* as if it were a special kind of *entity*: we have said that propositions are sentence-meanings, different from sentences or clauses. They are what is common to synonymous sentences, just as the number four is what is common to the Beatles, the Evangelists and the John Coltrane Quartet, along with every other four-membered set or collection. As we will see, it is very natural and useful to speak as if there really are these abstract entities, namely propositions, just as there are numbers. Later, in Chapters 5 and 6, we will consider reasons for scepticism about the idea.

● CONTEXT-DEPENDENCE

Now that we have declared that a proposition is the meaning of a sentence, we have to take it back slightly. Consider the sentence:

I am the father of Julius Caesar.

What is the meaning of this sentence? If propositions, as we are assuming, are the meanings of sentences, then *it does not have a complete meaning*. For the same sentence expresses different propositions depending on *who utters it* (and when, since there was a time when the man who became the father of Julius Caesar was not

yet his father). The word ‘I’ picks out or refers to different persons depending on who utters it. The sentence expresses a certain proposition if uttered by Julius Caesar’s dad, another one if uttered by Groucho Marx. There are many words like this – words that refer to different things depending on time, place, identity of speaker or hearer and other facts, concerning what we call the **context of utterance**. Further examples:

<i>here</i>	refers to place of utterance
<i>now</i>	refers to time of utterance
<i>you</i>	refers to person addressed by the speaker
<i>this, that</i>	refers to object indicated by the speaker

These are called **indexicals** (or ‘deictic’ expressions). They are simple and obvious examples of context-dependent expressions.

Somewhat less obvious sources of context-dependence include the tenses of verbs. Consider:

Octavian is Emperor of the Roman Empire.

The indexicality resides in the present tense of the verb ‘is’. This sentence would have expressed a false proposition before Octavian (Augustus Caesar) became Emperor in 27 BC, but a different, true one for a while after that (until his death in 14 BC, when the sentence reverted to expressing a false proposition). Parallel remarks go for other tenses, such as the future tense of ‘is’ – ‘will be’ – and the past – ‘was’. Since almost everything we say includes a tensed verb, almost everything we say is context-dependent at least with respect to time of utterance.

Terms such as ‘this’ and ‘that’ constitute a special class of indexicals called **demonstratives**: they often require an accompanying pointing gesture or suchlike in order to pick out a referent. It is standard to call the accompanying gesture or other device a **demonstration**.

For these reasons, it is really more accurate to say that propositions are features of *utterances* of sentences (actual or possible); better still, a proposition is determined by a sentence *with respect to*, or *at*, a context. A context, we will say, is a set containing at least the time of utterance, place of utterance, identity of speaker and audience, and objects indicated by demonstrations, if any. Thus associated with a sentence is a rule, or function, that determines what proposition, if any, the sentence would express at a given context (equally, the **statement** that would be made by uttering it at a given context). The picture is like this:

Sentence + Context \rightarrow Proposition

In fact, for the time being, we are mostly going to *ignore* all forms of context-dependence. We will pretend that declarative sentences always express the same proposition. We will return to this issue much later, in Chapters 8, 9 and 13.

● THE ROLES OF PROPOSITIONS

A proposition, we are assuming, is neither animal, mineral nor vegetable. It is not something that might be inspected with a microscope or a telescope. Like the number two, it is not a material object at all; it is an *abstract entity*. (Nor are they *mental entities*, as will emerge presently.) But we can characterise propositions in terms of certain roles they play and relations in which they stand. Similarly, we might not be able to say what the number two is ‘in itself’, or point to it, but we can say that it follows one, precedes three, is the number of ears belonging to King Charles, and so on. We have just said that the proposition that snow is white is the meaning of the three sentences above. That is one role of propositions: to be the meaning of a sentence (at a context of utterance). There are two more.

The second role of propositions concerns what Russell called the **propositional attitudes**. Consider John, Pierre and Hans. They speak respectively only English, only French, only German. But they all believe that snow is white.

We have:

- (8) John believes that snow is white.
- (9) Pierre believes that snow is white.
- (10) Hans believes that snow is white.

Intuitively, John, Pierre and Hans *believe the same thing*. That is, there is at least one thing that John, Pierre and Hans all believe. That thing is the proposition that snow is white. We shall regard the that-clause *that snow is white* simply as a singular term which stands for the proposition.

Consider the inference from (8) to (10) above to:

- (11) There is something that is believed by John, Pierre and Hans.

The reasoning from (8)–(10) to (11) certainly seems valid. If so, *then our normal way of reasoning about beliefs commits us to the existence of propositions*. And the reasoning seems to show that propositions cannot be mental entities: whereas the self-same proposition is common to all three believers, mental entities – an emotion, a dream, a pain – cannot be in several minds at once. Each mental entity can only be in one mind.

Propositions, then, are the objects of belief. To believe is to stand in a certain relation to a proposition. Belief is a propositional attitude: an attitude towards a proposition. There are other propositional attitudes: one may *believe* that the fish is fresh, but one may also *doubt* that the fish is fresh, *wonder whether* the fish is fresh, *hope* that the fish is fresh, and so on.

The third role that propositions play is that of truth-vehicles. Here it is useful to delve back again into context-relativity for a moment. Consider the following exchange:

Phocas: *I am the rightful Emperor of the Roman Empire.*

Maurius: *I am the rightful Emperor of the Roman Empire.*

Phocas and Maurius, of course, disagree. They correctly take it for granted that there is only one rightful Emperor of the Roman Empire. But they utter the very same sentence (two tokens of the same type). The semantical difference is that Phocas says that Phocas is Emperor, but Maurius says that Maurius is. Phocas implicitly denies what Maurius asserts, and Maurius implicitly denies what Phocas asserts. In fact, if Gibbon is to be believed, Maurius speaks truly, Phocas falsely. They use the *same* sentence but express *different* propositions. One and the same thing cannot be both true and not true (similarly, the same light cannot be both on and not on at a given moment). Thus this thing cannot be the sentence. The thing that is true is the proposition that Maurius is the rightful Emperor of Rome.

Propositions, then, are:

- (i) the meanings or contents of sentences (in contexts of utterance);
- (ii) the objects of propositional attitudes;
- (iii) the vehicles of truth and falsity (the things that can directly be true or false).

● COMPOSITIONALITY, STRUCTURE AND UNDERSTANDING

It is plausible that to understand a sentence is to *know what it means*. In view of the foregoing discussion, we can take this quite literally: since what a sentence means is the proposition it expresses, to understand a sentence is to *know which proposition it expresses*. But if you think about it, merely knowing what a sentence means is not quite sufficient for *understanding* it. If someone reliably tells me that a certain sentence of Urdu means that snow is white, I might thereby come to know that that sentence means that snow is white; but it seems wrong to say that I thereby come to understand the sentence. In order to understand a sentence I must also *know the meanings of the individual words and grasp its meaning on the basis of how it is put together*. Consider 'Snow is white'. What makes it right to say that I understand this sentence is that I know the meaning of 'snow', 'is' and 'white', and I understand the significance of putting those words together in that way.

So, still abstracting from the context of utterance, we can formulate and set off for emphasis:

The principle of compositionality: the meaning of a sentence is determined by

- (i) the *meanings of the words* it comprises; and
- (ii) the *semantic significance* of the *grammatical structure* of the sentence.

The first requirement is relatively transparent, but we must also stress the importance of the second requirement. It implies, for example, that merely having an

English–Burmese dictionary would not enable a Burmese speaker to understand sentences of English. To take an obvious illustration, the same words constitute ‘The dog bit the baby’ and ‘The baby bit the dog’, but the *order* of the words makes all the difference to the meaning.

We can dig a bit deeper. There are countless possible sentences of your language that you have never heard, spoken or read, but which you would readily understand if you did hear or read them. Similarly, one is endlessly creative; one’s ability to produce novel sentences is amazing but seldom remarked because it is so commonplace. How is that possible? Answer: because you know the meanings of the words they comprise, and you know the semantic significance of the syntactic structure of sentences.

In fact, even though each of us alas has a finite brain, and knows only finitely many words and grammatical principles, there are infinitely many sentences that this finite knowledge enables us potentially to understand. A trivial illustration: competent speakers of English can understand ‘He is her father’, ‘He is her father’s father’, ‘He is her father’s father’s father’ and so on, without upper bound. Of course, if it gets too long, then we might get confused or fall asleep before understanding it; the point is that the understanding one has of ‘father’ is sufficient, *in principle*, to determine the meaning of any of these sentences. This behaviour of ‘father’ is known as its being **recursive** (its being ‘iterative’): it is a basic example of the sort of thing that underlies the capacity for genuine creativity, of the capacity to comprehend novel sentences. It makes a potentially infinite capacity out of finite means. Insofar as we are like digital computers, the finitude of our actual capacity is due merely to the hardware, not to intrinsic limitations of the program, the software.

This capacity is often thought to mark the difference between human language and language-like behaviour of parrots, gorillas, dolphins etc. – they may use or respond appropriately to an impressive array of individual words or signs, and may even use them in combination to approximate rudimentary sentences, but it is controversial whether they show any evidence of genuine recursion. (On the other hand, there are movements afoot that seek to deny that compositionality is a feature of *all* human languages; Daniel Everett (2008) claims that it is not a feature of the language of the Pirahã of the Amazon basin, which is thought by some to be evidence that compositional structure is not hardwired as part of the human genetic endowment in the way that Noam Chomsky has famously argued.) But we have no need to take a stand on this; recursion is integral to the sorts of language we’ll be concerned with, and the exploits and potentials of other animals do not bear on the concerns of this book.

We leave this ahistorical introduction with a famous quote from Gottlob Frege, in his essay ‘Compound Thoughts’ of 1923 (part of the extended essay ‘Logical Investigations’; perhaps the first to enunciate the principle clearly was Alexander von Humboldt, in the first half the nineteenth century):

It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a human

being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thought corresponding to the parts of a sentence, so that the structure of the sentence serves as an image of the structure of the thought.

(Frege 1984, p. 390)

1

• naïve semantics and the language of logic

Language is an enormously complex phenomenon. As with many complex phenomena, it would be pedagogically extremely hard, in one fell swoop, to begin with a complicated theory covering all its many aspects. Compare physics, in which one studies a model of a ball rolling down a plane – ignoring friction, air pressure and resistance, imperfections in the ball and in the surface of the plane, and so on. We can learn a lot from the model, and think profitably about its most important features, without forgetting that the actual phenomenon is more complicated.

We will thus begin by considering a simple theory of language, one grounded in common-sense ideas of how language functions: **naïve semantics**. Later, one can adjust the theory or start over from a more informed perspective. It is not perhaps a theory which was explicitly held by anyone, and many philosophers and linguists hold that naïve semantics is almost *completely wrong*. But if so, it is wrong in something like the way that Newton's classical physics was: it is a good start and is intuitively satisfying in many respects. Further, in order to see why a different theory is needed, it's useful to see where it breaks down. This, then, will provide a basis from which to consider the more elaborate Frege–Russell outlook, which might be called *classical semantics* or the *classical theory of meaning*. They are the subjects of Chapters 2 and 3.

This chapter is a bit dry as well as philosophically comparatively barren, but it will serve to introduce some terminology, to introduce some notions and lingo that are pretty sure to remain standing in what comes after. It is not long.

• NAÏVE THEORY: SINGULAR TERMS, PREDICATES AND REFERENCE

A sentence is made of words. Words fall into different grammatical types or classes – *syntactical categories*. To these categories correspond different *semantical*

categories – categories of meaning. According to the classifications of traditional grammar, these include proper names, nouns, pronouns, verbs, adjectives, definite and indefinite articles, adverbs, prepositions, quantifiers and more. For the purposes of the philosophy of language, many of these distinctions don't matter; we'll carve up language in slightly different and in sometimes cruder ways – ways, primarily, that in the first instance directly affect the *truth-conditions* of statements. Also, the individual words are sometimes best treated not as semantically significant parts in themselves but only as parts of parts that do have meaning (as 'syncategorematic'). What all this means will be much easier to see once we get going.

We first introduce and explain the notion of an **atomic sentence**, such as 'Jane smokes' or 'The cat is on the mat' (also called elementary sentences).

SINGULAR TERMS

Consider:

- (1) Mars is red.
- (2) Mars orbits the sun.

Both (1) and (2) contain the name 'Mars'. 'Mars' is a name, a proper name, of Mars. It *stands for* it, *names* it, *picks it out*, *denotes* it, *designates* it. According to a decision announced in the Introduction, we say that it *refers* to it; Mars, the actual planet Mars with all its red dust, is its *referent* – i.e. the thing it refers to. It is customary to call 'Mars' a **singular term**. We will also formulate naïve principle 2 (naïve principle 1 will be introduced shortly):

(NP2) The meaning of a singular term is its referent.

We are not going to try to give a precise definition of 'singular term'. It is too hard. But our intuitive classification is sufficiently reliable: we are thinking simply of words whose *role*, whose *function*, is to stand for a single object, a certain individual – a person, a city, a planet (so all these things are *objects*, in an extended but philosophically standard sense of 'object'). 'Dog', by contrast, is not a singular term because it refers to many dogs.

Thus 'the sun', as it occurs in (2), is also a singular term. It refers to the sun. Here are more singular terms:

- (3) Jupiter
- (4) Lionel Messi's mother
- (5) the river that runs through Prague
- (6) the fastest mammal

(3) is a proper name, but (4) and (5) are not (though (4) and (5) *contain* proper names or titles as parts). As you can see, singular terms may be *simple* (containing no expressions as parts), as in the case of (3), or *complex*, as in the case of (5).

PREDICATES I: SYNTAX

If we remove the name ‘Mars’ from (1) and (2), we get:

(7) ____ is red.

(8) ____ orbits the sun.

These, in the logical sense of the word, are **predicates**. In general: *the result of removing a singular term from a sentence is a predicate*. This is a point of *syntax*, for we expressed it without speaking of the *semantics* of predicates, of their meaning.

In writing down (7) and (8) we used underlining to indicate blanks or gaps – the places vacated by the singular terms we removed. It is convenient to refine this practice, using Greek letters to indicate the gaps:

(9) α is red.

(10) β orbits the sun.

The Greek letters do not *mean* anything. Nor are they *variables*, either (as we use in logic to express quantification, or in algebra to speak of numbers in general, as in ‘ $2(x + y) = 2x + 2y$ ’). They are just there to mark the gaps, the places in predicates where names can be inserted. We call this procedure *predicate extraction*.

We can attach any singular term to a predicate such as (9) or (10), and the result is a sentence. In particular we can make a sentence in this way by replacing the Greek letter with a singular term. Thus we can attach (4) to (10), yielding the sentence:

(11) Lionel Messi’s mother orbits the sun.

It’s not likely that anyone would ever say this, but nevertheless there is nothing grammatically wrong with it as a sentence. It has a truth-condition, as we say.

The predicate (10), you may observe, contains a singular term. The sentence (2), from which we derived (10) by deleting a singular term, contains two singular terms, not just one. If we now delete the remaining singular term from (10), inserting another Greek letter in the vacated space, we extract:

(12) α orbits β .

This too is a predicate. But unlike (9) and (10), which are **one-place predicates**, or monadic predicates, (12) is a **two-place**, or binary predicate.

To construct another sentence from (12), we can replace both Greek letters with singular terms (either different ones or the same one used twice). There are also three-place predicates, such as:

(13) α gave β to χ .

There are in principle predicates of n places for *any* finite n , no matter how large.

Unlike (10), no further predicates can be extracted from (12) or (13) by removing singular terms. (12) and (13) are what we shall call **pure predicates**, by which we mean that not only are they without singular terms as parts, they also do not contain sentential connectives such as ‘and’ and ‘or’, or quantifier-words such as ‘something’ or ‘everything’. We shall consider sentential connectives and quantifiers below. We will also set aside adverbs such as ‘quickly’ (although the presence of an adverb does not disqualify a predicate from being a pure predicate).

We adopt a rule governing our use of Greek letters: when replacing Greek letters with names (or, later, with variables) to form a sentence, always replace *every* occurrence of each Greek letter with the same name (or variable). For example, we regard the following as *different* predicates:

α killed β .

α killed α .

Thus ‘Jones killed Jones’ is correctly obtained from either of these predicates, but ‘Jones killed Smith’ cannot be obtained correctly from the second one. This reflects the fact that the concept of *suicide* could be defined in terms of the concept of *killing* (one dies by suicide just in case one kills oneself), but that of killing could not be defined in terms of suicide.

Such sentences as (1), (2) and (11) are the simplest sentences: each is constructed from a pure predicate and the requisite number of singular terms, and contains nothing else, no other kind of expression. Indeed, any such sentence contains exactly one (pure) predicate. Sentences of this kind are called **atomic sentences**.

An **atomic sentence** is a sentence comprising nothing besides one n-place pure predicate and n singular terms.

WHAT ABOUT VERBS?

In traditional grammar, it is customary to say that every sentence must contain a *verb*. But the notion of a verb is of little use to us. In school grammar we say that ‘is’, as it occurs in (1), for example, is a verb. But what does that mean? What is a verb? The word ‘is’, unlike ‘orbits’, is not literally an ‘action word’. There is no *action* indicated in ‘Snow is white’. What is clear is that a sentence cannot be formed simply from two singular terms, as in:

(14) Mars Lionel Messi’s mother

Nor from predicates alone, as in:

(15) is red orbits is red

The words in (14) and (15), so to speak, don't stick together. Rather than the grammarian's claim about verbs, we repeat that atomic sentences are all and only those expressions that can be formed by correctly replacing all the Greek letters in a pure predicate with singular terms.

● PREDICATES II: SEMANTICS

The intuitive idea of a singular term is that of an expression that stands for something: its being meaningful is that it refers to an object, a particular individual. Our intuitive ideas about the meanings of predicates are less clear and more varied. But the following seems clear: whereas the meaning of a singular term – an object – is *what we say things about*, the meaning of a predicate is *what we say about* an object. This fits with what we said about the syntax of predicates: syntactically speaking, a (one-place) predicate is what is left over from deleting a singular term from a sentence; semantically speaking, the meaning of a (one-place) predicate can be significantly ascribed to any object.

A note aside. In the sentence 'Mars is red', which comprises a singular term and a predicate, 'red' is not a singular term like 'Mars'. Its semantical job is not to stand for an object; its job is not to refer to an individual. Rather, combined with 'is' to form a predicate, its role in this sentence, its semantical job, is to tell us the particular way Mars is being said to be by means of the predicate. It is a **general term**. Other general terms are 'dog', 'hungry', 'barking' (representing respectively nouns, adjectives and verbs). But mostly we will speak not of general terms but of the predicates in which they figure.

We'll say more about the meanings of predicates. Continue with (1) above: 'Mars is red'. Whereas 'Mars' refers to Mars, ' α is red' *applies* to Mars, since Mars is red. Another way to put the same idea is that the predicate is *true of* the object Mars. Yet another way is that the object *satisfies* the predicate.

In virtue of what does a predicate apply to an object? Not in virtue of another *object*. Although 'Mars is red' is true, the meaning of ' α is red' is not the planet Mars. If it were, then 'Mars' and ' α is red' would be synonymous, which they are not, and there would be no semantical difference between 'Mars is red' and 'is red is red' or 'Mars Mars', which obviously there is. Whereas Mars is the only thing named by 'Mars', ' α is red' applies to lots of other things besides Mars. And some pairs of predicates have their *different* meanings despite applying to just the *same* things; ' α is a continent bigger than Africa' and ' α is a continent that contains China' both apply to exactly one thing, namely Asia. So applying to a thing must be different from naming a thing: singular terms and predicates have different sorts of meanings, different semantical roles.

One candidate for the meaning of a predicate such as ' α is red' is an *idea* in the minds of users of the term; John Locke held something like this view in the seventeenth century. But, for various reasons, in the modern era the idea 'idea' has not proven popular. For one thing, such 'ideas' cannot be mental pictures or images,

because many predicates, even simple ones, stand for things for which it is implausible to say we have mental images, such as ' α is abstruse'. And in cases where it is not so implausible, your mental image and mine can be rather different – your idea of a dog may be of a French Bulldog, mine of an Irish Wolfhound – yet the predicate means the same in your mouth and mine, in the sense that the criterion for objects falling under the predicate ' α is a dog' doesn't vary depending on which dog-image we employ, French Bulldog or Irish Wolfhound. This is connected with the observation that whereas ideas are necessarily private and subjective, what we communicate, the meanings of our words, is in a certain sense objective and public. *Singular terms* may be connected in our minds with subjective and private ideas – for example, I might have particular sentimental ideas of Paris – but they are also connected with public objects as their referents – like the city of Paris itself, with its boulevards and cafés. Likewise with predicates; this way, sentences can express meanings with public parts corresponding to singular terms and predicates, so the fact of communication is no mystery.

A slightly better candidate, one that does not fall foul of what might be called the 'requirement of publicity', is the *set of red things*. More generally, the idea is that the meaning of a one-place predicate is the set of things that satisfy the predicate. In logic, semantics or the theory of meaning this is called the **extension** of a one-place predicate. Thus we might suppose that:

(NP3*) The meaning of a predicate is its extension.

The asterisk is there because we are not going to use this idea. More powerful reasons for rejecting (NP3*) will emerge later, but two simple reasons that are sufficient for now are as follows: (1) as already mentioned, many predicates that are satisfied by exactly the same set of objects have different meanings; in addition to the example given above, there are pairs of predicates such as 'is a round square' and 'is a human being taller than the Eiffel Tower', that have the same extension, namely the empty or null set, but their meanings differ; (2) the extension of the average predicate is always changing (every moment, some things become red, others cease to be red); yet it seems wrong to say that the *meaning* of the predicate thereby changes.

Much better is this:

(NP3) The meaning of a (one-place) predicate is the property for which it stands. The meaning of a (two-place) predicate is the relation for which it stands – similarly for three-place predicates, etc.

The property (or attribute, or quality) that 'is red' stands for is *redness*. Is it plausible to say that if 'is red' came to stand for a different property, then its meaning will have changed? Yes. Is it plausible to say that one understands 'is red' just in case one knows that it stands for redness? Yes. Is it plausible to say that properties are in the relevant sense public rather than private? Yes.

Note that our use of the word 'property' is much more liberal than the ordinary one. Ordinarily, we use it to mean such scientifically interesting features as *hardness*,

flexibility and the like, whereas we intend the meaning of any one-place predicate, such as ‘ α used to go to bed early’. A property is any *way that a thing can be*.

Consider now two-place predicates, such as ‘ α orbits β ’. Such predicates stand for *relations* (two-place, or *binary* relations, in this case). Relations, like properties, are *universals*. So we could say: the meaning of a predicate is a universal; for example, both whiteness and love are universals, but the latter is a relation, the meaning of ‘ α loves β ’.

Thus according to the naïve theory, properties and relations – universals – are simply the referents of predicates. So the meaning of a singular term *or* a predicate is its referent – it is either an object, property or relation. Now we have not yet considered all types of expression, but we have said enough perhaps to motivate:

(NP1) *The fundamental principle of naïve semantics*: the meaning of every expression is its referent.

The idea is that the meaning of an expression is what it stands for: *meaning = reference*.

● TRUTH AND MEANING FOR ATOMIC SENTENCES

(A) ATOMIC PROPOSITIONS

Remember from the Introduction that the meaning of a sentence is the proposition it expresses (in a context of utterance). So that idea, along with NP1, entails that propositions are the referents of sentences. This is a little weird. But we do speak (naïvely!) of *facts* as what true sentences stand for; and if facts are just true propositions, then maybe NP1 is not so weird after all.

Since atomic sentences are sentences, their meanings are propositions. We call these **atomic propositions** (we could call them ‘elementary propositions’). According to the principle of compositionality, the meaning of a sentence is determined by the meanings of its parts, together with its structure. And what, according to naïve semantics, do the meanings of the parts of an atomic sentence have to do with the meaning of the sentence? Simple: the proposition expressed by an atomic sentence – its meaning – is actually *composed* of the meanings of the parts of the sentence. The proposition is a *composite object* whose parts are the meanings of the sentence. Thus the proposition expressed by an atomic sentence is composed of, made up of, the referents of the expressions occurring in the sentence:

(NP4) *Naïve semantics for atomic sentences*: the proposition expressed by an atomic sentence containing n singular terms is composed of the referent of its predicate and the n referents of those singular terms.

We can think of an atomic sentence as being like a picture or diagram of its meaning. The proposition expressed by ‘Mars orbits the sun’ can be shown in diagram form, as in Figure 1.1. The structure is an abstract sort of structure – the mode of

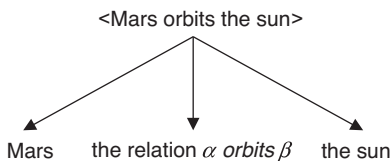


Figure 1.1 The proposition expressed by 'Mars orbits the sun'

composition is not physical – but otherwise it is very like a physical structure. The predicate is something like a pod, and the referents of the singular terms are the peas (there are other possibilities as to the nature of propositions, but here we shall not explore alternatives to this 'structural' conception).

(B) TRUTH FOR ATOMIC PROPOSITIONS

Remember that propositions are the vehicles of truth. The things to which the predicates 'is true' and 'is false' immediately attach are not sentences but the propositions they express.

We can now define what it is for an atomic proposition to be true. For simplicity, we continue to confine our attention to atomic propositions containing only properties or binary relations. We call these *one-place* atomic propositions and *two-place* atomic propositions.

Common sense seems to approve the idea that truth is in *correspondence with reality*. The simplest way to explain this is as follows.

- (a) The proposition *that Mars is red* is true if and only if the object Mars *possesses* the property redness. That is, it is true if and only if the object has that property – *exemplifies it, instantiates it*.
- (b) The proposition *that Mars orbits the sun* is true if and only if those two things stand in the relation denoted by ' α orbits β ': that is, it is true if and only if Mars bears that relation to the sun – stands in that relation to it.

Thus in general:

(NP5) Naïve definition of truth for atomic propositions: An atomic proposition is *true* if and only if (i) it is a one-place proposition comprising an object o and property P , and o possesses P ; or (ii) it is a two-place proposition comprising an object o_1 , a relation R and an object o_2 , and o_1 bears R to o_2 .¹

Note finally that none of this actually prevents us from speaking of *sentences*, or *utterances* or *statements*, as being true or false. A sentence, utterance or statement is rightly called 'true' if and only if the proposition it expresses is true. According to this way of thinking, to speak of a sentence as being true is simply elliptical for speaking of a proposition as true. Similarly with beliefs: a belief is true just in case the proposition believed is true.

• LOGICAL SYNTAX AND LOGICAL OPERATORS

So far, we have been speaking vaguely of ‘language’, and using English to illustrate the points we wish to make. But the idea is that *all* languages feature atomic sentences, hence singular terms and predicates. Indeed, the naïve idea is that all adequate languages *must* have such devices, because the *world* is made up of objects having properties and objects standing in relations; a language without singular terms and predicates could not talk directly about the world, could not represent it, because it could not reflect its structure. As the great logician Alfred Tarski once put it, the idea of the semantical categories cuts so deeply into our intuitions that it is difficult to imagine the possibility of an alternative. Thus we have been using English merely to *illustrate* points whose import is *fully general*.² So although we will stick to English, you should bear in mind that we are trying to theorise about language in general. English is not our actual subject-matter.

This section is about non-atomic sentences, or ‘molecular’ sentences. The trouble is that once we move beyond atomic sentences, English is complicated and confusing in ways that, for our purposes, are best avoided as much as possible. For now, we will restrict our attention to a simplified version of English, and to just a few devices in English, ones corresponding to the **sentence-connectives** and **quantifiers** that you may have learned in a logic course. But we won’t assume that you have had such a course, and our use of them will be quite simple.

SENTENCE-CONNECTIVES

We will make regular use of the expressions ‘not’, ‘or’ and ‘and’. We will assume that these are **truth-functional**, by which we mean that for any two sentences p and q :

Negation: ‘Not p ’ is true if and only if p is false.

Conjunction: ‘ p & q ’ is true if both p and q are true; false otherwise.

(Inclusive) Disjunction: ‘ p or q ’ is false if both p and q are false; true otherwise.

For exclusive disjunction, write ‘ $(p \text{ or } q) \text{ \& not } (p \text{ \& } q)$ ’. Negation in English is expressed in a variety of ways. For simplicity, we take advantage of the fact that it can always be expressed in the form ‘It is not the case that’. Since that is a pain to write out, we just write ‘Not: ____’. For example, for ‘It is not the case that Jupiter orbits Neptune’ we write:

Not: Jupiter orbits Neptune.

It is more controversial whether the conditional ‘if–then’ of English is really truth-functional. For now, we will ignore the question and use the arrow ‘ \rightarrow ’ of formal logic. So instead of ‘If Charles is tall then Andrew is tall’, we write:

Charles is tall \rightarrow Andrew is tall.

And we define:

Conditionality: ' $p \rightarrow q$ ' is false if p is true and q is false; true otherwise.

Thus if Charles is tall, then the conditional is true if Andrew is tall and false if Andrew is not tall. If Charles is not tall, the conditional is true. It is also useful to have 'iff', that is, 'if and only if':

Biconditionality: ' p iff q ' is true if ' $p \rightarrow q$ ' and ' $q \rightarrow p$ ' are both true; false otherwise. (Or: ' p iff q ' is true if p and q have the same truth-value; false otherwise.)

QUANTIFIERS

The expression of generality is very complicated and messy in English, but we can for the most part ignore the complexity. Consider a statement that does not involve generality:

If Fido is a retriever then Fido swims.

A person saying this would presumably be relying on a certain background generalisation, which can be expressed in any number of ways, such as:

All retrievers swim.

Any retriever swims.

Every retriever swims.

In this book, we proceed as follows. First, we can express the above somewhat awkwardly as:

For every thing, if it is a retriever then it swims.

We think of this as comprising a **universal quantifier** – 'For every thing' – and an open sentence or 'matrix':

if it is a retriever then it swims.

Then in place of the pronoun 'it', we introduce **variables**:

if x is a retriever then x swims.

According to our policy with the conditional, we can rewrite it as:

x is a retriever $\rightarrow x$ swims.

Then we restore the quantifier with the variable written in the place of ‘thing’ and put in a pair of parentheses for clarity:

For every x (x is a retriever $\rightarrow x$ swims).

The above sentence is a universal quantification – the quantifier ‘binds’ the variable x . We could just as well have written ‘For *each* x (x is a retriever $\rightarrow x$ swims)’, or ‘For *all* x (x is a retriever $\rightarrow x$ swims)’. Same thing. The important thing is that the sentence:

For every x ——

is true as long as the sentence written in place of ‘——’ is true for every choice of x and not otherwise. Thus the ‘open sentence’:

x is a retriever $\rightarrow x$ swims.

is true no matter what object x is. ‘For every x (x is a retriever $\rightarrow x$ swims)’ is a ‘closed’ sentence, the universal closure of the ‘open’ sentence ‘ x is a retriever $\rightarrow x$ swims’.

We will also use the **existential quantifier**, which we write as:

There is an x such that ——

As in:

There is an x such that (x is an albino & x is a tiger).

This sentence is to be understood as true if there is least one albino tiger and false otherwise. Equally, we could say that *there exists an x such that* x is an albino & x is a tiger, or *for some x* , x is an albino & x is a tiger. Same thing.

Occasionally we’ll need more variables – ‘ y ’ and ‘ z ’ in addition to ‘ x ’. For example:

For every x , there is a y such that (x admires y).

That’s our way of saying ‘Everyone admires someone’. ‘Some girl is admired by every boy’ goes as:

There is an x such that (x is a girl & for every y (y is a boy $\rightarrow y$ admires x)).

This sort of example shows why we use variables instead of sticking with the ordinary pronouns: pronouns are hard to keep track of, and sometimes ambiguous, when there is more than one quantifier involved. And for those not averse to them, ‘ \forall ’ and ‘ \exists ’ of formal logic can be written instead of our ‘For every...’ and ‘There is a... such that’, along with the variables.

Words such as ‘all’, ‘any’, ‘every’, ‘some’, ‘there are’ and ‘there exists’ are examples of natural-language quantifiers. Additional natural-language quantifiers are words like

‘more’, ‘none’, ‘seven’, ‘at least two’, ‘half’ and so on – any word or group of words which expresses quantity, of how much or how many of a thing or kind of thing is being talked about. The standard view is that quantifiers are not singular terms, and nor are they predicates on the same level as ‘ α is green’. Why that is and what category they do come from is somewhat contentious, and we’ll hear more on the subject in Chapters 2, 3, and 4. But note that from the steadfastly naïve point of view, they, along with the sentential connectives, are referring expressions like every other expression of language, even if they are not singular terms or predicates on the same level as ordinary predicates. (Frege himself, the subject of the next chapter, thought that quantifiers referred to higher-order attributes – attributes of attributes.)

● GENERALISING AT THE LEVEL OF SINGULAR TERMS AND PREDICATES

Occasionally, we will want to reason purely schematically, using *uninterpreted* singular terms and predicates. That is, we might want to say something that holds for ‘Socrates taught Plato’ but in such a way that it holds for any sentence comprising two singular terms joined by any two-place predicate. So we’ll just write ‘ a ’, ‘ b ’ and so on for uninterpreted singular terms, ‘ Fa ’, ‘ Ga ’ for arbitrary one-place predicates, and ‘ $Ra\beta$ ’ for an arbitrary two-place predicate. Then we write, for example:

Fa

and

Rab

for atomic sentences (the more suggestive ‘ aRb ’ for the last one is out of fashion). This follows the standard practice in formal logic, which is itself modelled on mathematics.

● HISTORICAL NOTES

It’s improbable that any actual person explicitly held the naïve theory of semantics exactly as presented here. Among ancient philosophers – primarily the Greeks – Plato (especially in his dialogue the *Cratylus*), Aristotle and the later Stoics explored some of the above ideas. The medieval and scholastic philosophers, Indian philosophers and the so-called ‘modern philosophers’ of the seventeenth and eighteenth centuries – Locke, Leibniz, Hume and Berkeley – all wrestled at least occasionally with problems concerning language and meaning with something like the naïve theory presupposed. Through a variety of factors, not the least of which was the development of logic in order to make good on the idea that mathematics is ultimately just logic, the philosophy of language became more explicit and central to philosophy at the end of the

nineteenth century. In particular Gottlob Frege, in his early *Begriffsschrift*, of 1879, and Bertrand Russell, especially in his *Principles of Mathematics*, of 1903 (not to be confused with *Principia Mathematica* of 1910), subscribed to versions of the naïve theory; so did Ludwig Wittgenstein in his early, 1921 book *Tractatus Logico-Philosophicus*. An interesting precursor from our perspective is John Stuart Mill (1806–73); in his *A System of Logic* of 1843, Mill did not have the ideas that would propel Frege and Russell into greatness in logic – principally, he lacked the idea of quantification – but he articulated the key naïve idea that proper names have denotation but not connotation – that, as it would be put much later by Ruth Barcan Marcus (1921–2012), they are just ‘tags’ of their referents. As we’ll see later, Russell never really departed from that idea, and it has its staunch supporters today. It was Frege who reacted in later work – especially ‘On Sense and Reference’ of 1892 (in Frege 1997) – to the apparent shortcomings of the idea and came up with a new way of thinking.

● CHAPTER SUMMARY

Naïve semantics begins with atomic sentences and breaks them up into predicates and singular terms. Each atomic sentence is composed of one n -place pure predicate and n singular terms. The meaning of a singular term is its referent, the object for which it stands; the meaning of a predicate is the universal or attribute that is its reference, that for which it stands. Universals comprise properties, binary relations, tertiary relations, etc. corresponding to one-place predicates, two-place predicates, three-place predicates and so on.

The meaning of a sentence is the proposition for which the sentence stands. The meaning of an *atomic* sentence is a composite abstract object comprising the objects corresponding to its singular terms and the universal corresponding to its predicate. The truth-value of an atomic sentence is determined by its correspondence to reality or its lack of correspondence with reality: it is true if there is a fact composed as the sentence says, and false if there is no such fact.

To cope with sentences that are not atomic sentences, we add (1) the devices of sentential logic – ‘not’, ‘or’, ‘and’ and ‘if-then’ – and (2) quantifiers corresponding to ‘all’ and ‘some’. For simplicity, we understand the former as truth-functional, in terms of the truth-tables in classical logic.

● STUDY QUESTIONS

- 1 The Vltava = the river featured in *By Night under the Stone Bridge* by Leo Perutz. Does this fact represent a difficulty for naïve semantics? How is it connected with the distinction between simple singular terms and complex ones?
- 2 Take any newspaper or magazine article and practise (a) identifying plausible examples of atomic sentences, (b) identifying the predicate in such sentences and (c) writing the predicate with Greek letters and then forming new sentences with the predicate, ensuring correct replacement.

- 3 According to naïve semantics, what, if anything, is the syntactical or semantical difference between the expressions 'redness' and ' α is red'?
- 4 According to the naïve theory, 'Jane' refers to Jane, and 'smokes' refers to the property of being a smoker. Is it not possible, then, to merely refer to Jane, and refer to the property of being a smoker, by saying 'Jane, the property of being a smoker' – without expressing the proposition that Jane smokes? What more is required in order to express the proposition?

● PRIMARY READING

Frege, G. (1997) 'Begriffsschrift (1879): Selections,' in *The Frege Reader*.

Mill, J. S. (1963ff) *System of Logic, Ratiocinative and Inductive*, volumes 7–8 of the *Collected Works of John Stuart Mill*, Book I, Chapter II, 'Of Names'.

Russell, B. (1903) *The Principles of Mathematics*, Chapter IV, 'Proper Names, Adjectives and Verbs'.

● NOTES

- 1 We are leaving aside the worry about the inner structure of atomic propositions – the difference in 'order' whereby *Mary kissed John* differs from *John kissed Mary*. See Chapter 4.
- 2 This should not be confused with the concept of Universal Grammar first proposed by the linguist Noam Chomsky (e.g. Chomsky 1965). Chomsky's own view of the matter was that 'UG' is an empirical hypothesis about the structure of the human mind, not a metaphysical hypothesis about the very possibility of language. See Chapter 12.

2

• fregean semantics

• TWO PROBLEMS FOR NAÏVE SEMANTICS

The fundamental principle of naïve semantics is:

(NP1) The meaning of every expression is its referent.

In what follows, we concentrate on two corollaries of the principle, namely:

(NP2) The meaning of a singular term is its referent.

(NP3) The meaning of a (one-place) predicate is the property for which it stands. The meaning of a (two-place) predicate is the relation for which it stands (similarly for three-place predicates, etc.).

Gottlob Frege held something like the naïve theory in his early work. But in papers and notes written in the 1890s he formulated what are in effect serious objections to NP2 (and to NP3) and came up with a more complicated theory that purports to avoid them. The theory has a kind of overall symmetry that makes it very compelling, and its central components have been extremely influential, not only in the study of language but in the philosophy of mind and epistemology. In this chapter, we describe the problems in the naïve theory that motivated Frege's theory and then present a version of Frege's theory, or *Fregean semantics*, as we will say.¹

THE PROBLEM OF COGNITIVE VALUE

We begin with Frege's famous example. The Evening Star – also known as Hesperus – is the Morning Star – also known as Phosphorus. The bright celestial object that appears in the western sky just after sunset is the very same object as the bright

celestial object that appears in the eastern sky just before sunrise, namely the planet Venus. It is just one object appearing in different places at different times. Thus the following is true:

- (1) The Morning Star = The Evening Star.

According to the naïve theory, meaning is reference. In particular, the meaning of a singular term is its referent, namely the object for which it stands (NP2). Thus according to the naïve theory we have:

- (2) The meaning of 'The Morning Star' = the meaning of 'The Evening Star'.

Is this really correct? According to the principle of **compositionality**, the meaning of a sentence is determined by the meanings of its parts, together with the way it is constructed from those parts. It follows that if two sentences have the same structure, and each part at each location in one sentence has the same meaning as its corresponding part in the other, then the two sentences mean the same thing – that is, they express the same proposition (since the meaning of a sentence is a proposition). It follows, then, that the following sentences mean the same thing, express the same proposition:

- (3) The Morning Star is a planet.
(4) The Evening Star is a planet.

According to naïve theory, then:

- (5) The proposition expressed by 'The Morning Star is a planet' = the proposition expressed by 'The Evening Star is a planet'.

But this seems obviously incorrect. (3) and (4) do not seem to say exactly the same thing. So (5) is false: the propositions expressed by (3) and (4) are not the same. If there is any doubt about this, remember that a proposition is supposed to be the object of propositional attitudes such as belief. Thus if (5) were true, then the belief that the Morning Star is a planet would be the *very same belief* as the belief that the Evening Star is a planet. But clearly they are not: one could coherently believe (3) without believing (4). This could happen if one did not happen to know that the Evening Star is the Morning Star (in fact, there was a time when people did not know this) or if one positively denied that the Evening Star is the Morning Star. But if one could believe the one proposition but not the other, then they cannot be the same proposition. Likewise, if one kicks X but does not kick Y, then X and Y cannot be the same thing.

The matter emerges more dramatically if we consider the following:

- (6) The Morning Star = The Evening Star.
(7) The Morning Star = The Morning Star.

As Frege puts it, (7) is knowable *a priori*, without any astronomical investigation at all. It is a triviality that does not ‘extend our knowledge’. (6), on the other hand, is not a triviality, and is not knowable *a priori*; it required astronomical investigation to discover it, a discovery which *did* extend our knowledge. As Frege puts it, the **cognitive value** of (6) and (7) is not the same. Since the only difference between (6) and (7) is that one contains an occurrence of ‘the Morning Star’ where the other contains an occurrence of ‘the Evening Star’, the difference must be a difference in cognitive meaning between those two terms. Yet this cannot be a difference in reference; they refer to the same celestial item.

The examples have so far concerned *complex singular terms*, not simple singular terms – i.e. not, in the ordinary sense, *proper names*. Yet the same sort of thing arises in the case of proper names as well. For example, an ancient Babylonian name for the Morning Star is ‘Phosphorus’, and an ancient name for the Evening Star is ‘Hesperus’. Thus an ancient Babylonian might have known that (8) is true but not known that (9) is true:

- (8) Hesperus = Hesperus.
- (9) Hesperus = Phosphorus.

Likewise some people around you might not know that Mark Twain = Samuel Clemens.

THE PROBLEM OF EMPTY SINGULAR TERMS

Empty singular terms are singular terms with no referent. For example, towards the end of the nineteenth century, many astronomers were convinced by the French mathematician Le Verrier that there must be a small planet between the orbit of the planet Mercury and the sun, whose gravitational field caused the small perturbations observed in the orbit of Mercury. Le Verrier dubbed this purported planet Vulcan. Since Mercury is very hot, and it is close to the sun, Le Verrier, of course, would have believed the proposition expressed by:

- (10) Vulcan is hot.

Yet according to the naïve theory, *there is no such proposition*. For, by the beginning of the twentieth century, it became apparent there was no such planet; and the apparent perturbations in Mercury’s orbit can be explained by Einstein’s theory of relativity. Thus if there is no such planet, then ‘Vulcan’ has no referent, and therefore, according to the fundamental principle of naïve semantics (actually NP2), no meaning. But how could there fail to be such a proposition? It was *believed* that Vulcan is hot; how can someone believe when there is no proposition believed? A belief, surely, must have some particular proposition that is the *content* of the belief. It is as if someone were to *eat* when there was *nothing that he ate!*

If there is any doubt about this, consider this sentence, where ‘Nessie’ is a purported name of the Loch Ness Monster, which we assume does not exist:

(11) Nessie is hot.

‘Nessie’ is a singular term without a referent. According to the naïve theory, it is therefore meaningless. Therefore (10) and (11) must have exactly the same meaning: they are both atomic sentences that attach the meaning of the one-place predicate ‘*x* is hot’ to a meaningless singular term. But it seems clearly wrong to say that (10) and (11) mean the same thing.

• THE SENSE–REFERENCE DISTINCTION

These considerations motivated Frege to conclude that associated with a singular term, in addition to its *reference*, is something else, which he calls its *sense*. Singular terms with the same reference may have different senses – or, as we will put it more concisely, **co-referential** singular terms may have different senses.

That much, however, seems only to *name* the problem; what we want is a *theory* that predicts the phenomena we have noticed, and explains why they should arise. To do that, we have to answer the question: what is the sense of a singular term?

Frege characterises the sense of a singular term in two ways, only the first of which is prominent in the famous paper ‘On Sense and Reference’:

Sense I: the sense of a singular term is a *mode of presentation* of the referent.

Sense II: the sense of a singular term is a *rule for determining* its referent.

According to I, the sense is like a perspective on a thing; it is a way the object is presented to us – a way of thinking of the object. This has a strong overtone of Kant, according to whom there is no such thing as perceiving or thinking of an object *directly*: we must always grasp it in some *manner*, think of it or perceive in some *way*. According to II, the sense is a way in which we find a thing; it is a way that we *go to* the object instead of a way that it *comes to* us. Two senses that lead to the same object are like different routes to a summit, or different instructions for getting to the same place. But the two ideas can be brought together: one way in which an object can be presented to us is by its being the outcome of a search or procedure. *The sense presents the object as that which is determined by the rule.*

How is sense related to reference? It is tempting to represent the idea as in Figure 2.1 – as if all we have to do is to acknowledge that a singular term has two kinds or dimensions of meaning, namely sense and reference.

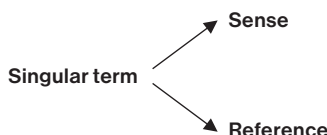


Figure 2.1 Sense and reference

This is not literally incorrect, but it fails to capture the relation that Frege envisages between sense and reference. A singular term has a referent *by virtue* of the sense it expresses: the term expresses a rule for picking out an object, and the object is the term's reference because the rule picks out that object.

A more accurate and informative diagram is Figure 2.2. Reference, the relation between word and thing, emerges as a complex relation: to *refer* to x is to *express a sense that determines* x . It is defined as that complex relation, just as *x is the maternal grandfather of y* is defined as *x is the father of the mother of y* . The curved line should not be regarded as indicating a relation that is distinct from the compound relation *x expresses a sense which determines y* .

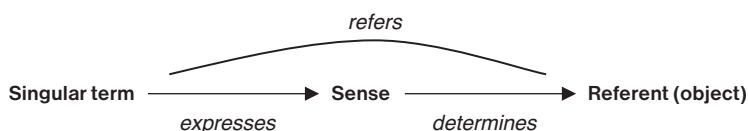


Figure 2.2 The relation of sense to reference

Now that we have distinguished sense from reference, what about *meaning*? Is the meaning of a singular term its sense, its reference, both or neither? English translations of Frege have confused this issue; what we are calling ‘reference’ (Frege: *Bedeutung*) has been translated both as ‘meaning’ and as ‘reference’ (also as ‘denotation’, ‘designation’ and ‘nominatum’). Yet in English it is more natural to use ‘meaning’ for what we are calling ‘sense’ (Frege: *Sinn*). Nothing really depends on this, however; it is only a matter of words. When discussing Frege, it is best just to use the terms ‘sense’ and ‘reference’ and avoid ‘meaning’.

Some translations use ‘meaning’ for *Bedeutung*, and some use ‘reference’ for *Bedeutung*. But luckily none use ‘meaning’ for *Sinn*; they always use ‘sense’. That makes it easy to avoid confusion when reading them: wherever you see the word ‘meaning’, think ‘reference’.

• THE DISTINCTION EXTENDED

Frege draws the sense–reference distinction for *all* expressions (we’ll begin to see why a bit later). For simplicity, we will confine the discussion to atomic sentences, hence to singular terms and predicates. Let us first consider sentences.

SENTENCES

Remember the principle of compositionality: the meaning of a sentence is determined by its parts, together with the way they are put together. (3) and (4), we said, express different propositions. But their corresponding *parts* all have the same referents. Thus the proposition expressed by a sentence is *not* determined by the referents of the parts of the sentence. The proposition expressed by a sentence, rather,

is determined by the *senses* of its parts (along with the way the sentence is put together). Frege calls the proposition expressed by a sentence a *thought*. The *sense* of a sentence is a thought. We will mostly stay with the word 'proposition': the sense of a sentence is a proposition. Do sentences have reference as well as sense? Yes, but we will come back to this point in a moment.

PREDICATES

Frege's actual view of predicates is a somewhat delicate matter. For now, we are going to work with a simplified scheme which respects important Fregean doctrines but is easier to work with. The *reference* of a predicate, we will say, is its **extension**. In the case of a one-place predicate such as 'is wise', this is simply the class (the set) of things that the predicate applies to – e.g. the class of wise things. The class of wise things is the class of things that the predicate is *true of*, or that **satisfy** the predicate. In the case of a relational (two-place) predicate, the extension can be understood as a set of ordered pairs. For example, the extension of the predicate ' α loves β ' would be the set of all ordered pairs $\langle x, y \rangle$ such that x loves y :

The reference of ' α loves β ' = $\{\langle \text{Victoria, David} \rangle, \langle \text{David, Victoria} \rangle, \langle \text{Charles, Camilla} \rangle, \langle \text{Camilla, Charles} \rangle \text{ and so on} \}$.

The *sense* of the predicate is a *rule*, or *criterion*, for whether or not an object or ordered pair goes into the extension of the predicate. The sense of ' α is happy', for example, is the criterion for whether something is happy; that is what we understand by the word 'happiness'. Similarly, the sense of ' α loves β ' is the criterion for love and is what we understand by it.

● COMPOSITIONALITY AGAIN; THE REFERENCE OF A SENTENCE

The sense of a singular term is a condition that an object must satisfy to be the referent of the singular term. The sense of a predicate is a condition that determines the extension of the predicate, its referent. In both cases, the sense is a *reference-determining condition*. Something similar holds in the case of sentences. Consider a simple atomic sentence:

Socrates is wise.

We said that the sense of a sentence is not determined by the referents of its parts. But, since the reference of a singular term is an object, and the reference of a predicate an extension, there is something that is determined by the referents of the parts: the *truth-value* of the sentence.

Think about it: suppose a certain object O is the referent of the singular term ' a ', then if O is a member of the extension of the predicate ' β is F ', then the sentence ' a is F ' must be true; if it is not, then the sentence must be false.

Thus the sentence is true if the referent of the singular term belongs to the extension of the predicate, and false if it does not. Furthermore, *if we replace the singular term with one with the same referent, the truth-value of the sentence will not change; similarly, it will not change if we replace the predicate with one with the same referent (same extension).*

Thus the truth-value of a sentence is determined by the referents of its parts. This leads Frege to say that the truth-value of a sentence is the referent of a sentence. This, in turn, enables us to bifurcate the principle of compositionality into two principles:

Compositionality of reference: the truth-value of a sentence is determined by the referents of its parts (along with the way the sentence is composed).

Compositionality of sense: the thought (proposition) expressed by a sentence is determined by the senses of its parts (along with the way the sentence is composed).

This enables us to say that the sense of *any* expression is a reference-determining condition. In particular: *the sense of a sentence is its truth-condition*. It is the condition whose satisfaction makes the sentence true.

In fact, there is a third principle of compositionality:

Compositionality of syntax: the sentence is determined by the syntactical properties of its parts and the way that the parts are combined.

The theory has a certain architectonic quality, a certain symmetrical interdependence of parts, which many people find compelling, as illustrated in Figure 2.3.

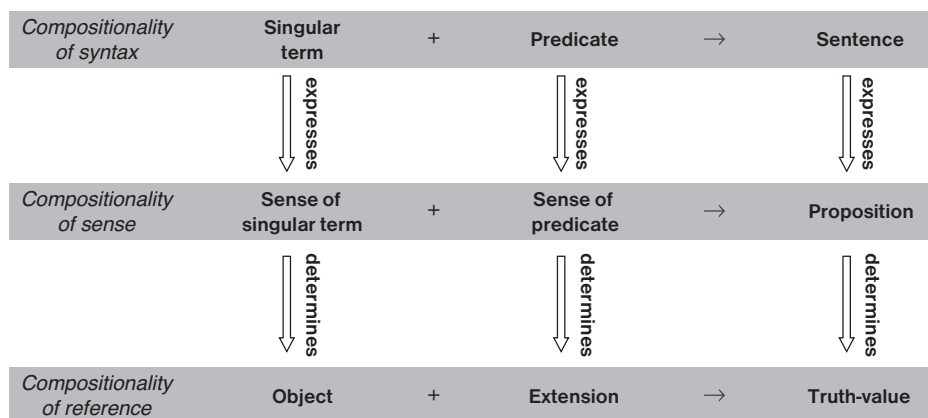


Figure 2.3 Fregean semantics for one-place atomic sentences

The idea is that when uttering a meaningful sentence, we express a proposition which is true or false depending on the referents of the parts of the sentence, which are themselves determined by the senses of the parts of the sentence.

Suppose, for example, that someone says 'Mikhail Gorbachev is dead'. The singular term 'Mikhail Gorbachev' expresses a *reference-determining condition* that determines a certain object, namely Gorbachev himself, the last premier of the Soviet Union. The predicate '*a* is dead' expresses the condition under which an object is a member of a certain extension, which is simply the set of dead things. These two items of sense combine to form the proposition *that Mikhail Gorbachev is dead* (the proposition expressed by 'Mikhail Gorbachev is dead'). This proposition is true if and only if Mikhail Gorbachev is a member of the set of dead things. Thus, as expected, the truth condition of 'Mikhail Gorbachev is dead' is that Mikhail Gorbachev is dead.

● APPLYING THE THEORY

How exactly does the distinction between sense and reference purport to solve the problems of cognitive value and empty singular terms?

Take the problem of cognitive value first. Pairs of co-referring singular terms, such as 'the Morning Star' and 'the Evening Star', may differ in sense. Thus (3) and (4) need not express the same thought, even if they must, given the truth of (1), have the same truth-value. Similarly, (6) and (7) need not, and do not, express the same thought. Since thoughts (propositions) are the objects of propositional attitudes, one might believe (7) without believing (6).

Now take the problem of empty singular terms. A singular term may express a sense that fails to pick out an object. For example, 'the largest natural number' or 'the present King of France' clearly expresses a reference-determining condition, but the condition fails to determine an object. A sense is also expressed by 'Vulcan': an object is the referent of that singular term just in case it is the unique planet orbiting the sun inside the orbit of Mercury. Thus there is a proposition expressed by (12); similarly for (13).

(12) Vulcan is hot.

(13) Nessie is hot.

So a sentence with reference-failure like (12) is meaningful – expresses a proposition – but is it *true* or *false*? The answer is neither. Since the sense of 'Vulcan' fails to determine an object, *no truth-value* is determined by the sense of (12). But the objection to the naïve theory was not that it fails to count such a sentence as (12) as having a truth-value but that it fails to count it as meaningful. Frege's theory does so. It expresses a proposition – the sense of the sentence – and thus can be meaningfully asserted: one that can be believed, doubted and so on. And we get the right result that (12) and (13) express different propositions.

Trouble: Unfortunately, Fregean semantics runs into trouble with respect to the closely related challenge presented by so-called **negative singular existentials**. Consider:

(14) Vulcan does not exist.

That sentence not only expresses a proposition: that proposition is true. Vulcan does not exist, and it's true that Vulcan does not exist. But according to Frege, the non-existence of a referent for the singular term shows that the sentence *lacks* a truth-value. The sentence instructs one to find Vulcan, and to see whether it satisfies the predicate '*a* does not exist'. But no such object is to be found.

Note that the problem does not arise for negative existentials in general. It arises only for negative *singular* existentials. For example, 'Dragons do not exist' is equivalent to 'For every *x*, *x* is not a dragon'. We can take each object (the Eiffel Tower, your left elbow and so on) and check it for whether it is a dragon; we will find of each object that it is not a dragon – no object satisfies the rule for being a dragon – thereby evaluating the original statement as true.

Frege did claim that the predicate '*a* exists' only misleadingly expresses existence; rather, existence is properly expressed by the quantifier 'There exists an *x* such that...' To assert existence in the general case – for example, 'Dolphins exist' – we write 'There is an *x* such that (*x* is a dolphin)'. In fact, in Frege's scheme, a quantifier expresses a *second-order* concept, a concept of concepts; in this case, we are saying of the concept *dolphin* that it satisfies the second-order concept of *being non-empty*, that there is something which fulfils it. To gain the effect of asserting existence of an individual object – for example, that the planet Jupiter exists – we write 'There is an *x* such that (*x* = Jupiter)'. The concept which is said to be fulfilled is the concept of being identical with Jupiter (or *being Jupiter*).

Thus it seems that the sentence attributing *nonexistence* to Vulcan should be written as:

(15) Not: There is an *x* such that (*x* = Vulcan).

But this doesn't work either. What we want is for (15) to say that the open sentence:

(16) *x* = Vulcan

comes out false for every *x* – thus ensuring that 'There is an *x* such that *x* = Vulcan' is false, hence that (15) is true. But according to Frege's theory, (16) comes out as lacking a truth-value, as neither-true-nor-false, for every *x*, on account of the presence of 'Vulcan'. 'The Eiffel Tower = Vulcan', 'My left elbow = Vulcan' and so on – all of these will be neither-true-nor-false. Thus 'There is an *x* such that *x* = Vulcan' is neither-true-nor-false, and likewise (15) is neither-true-nor-false. It is still the wrong result; the 'semantic motor' cannot start without some objects fed in.

This strategy of explaining existence with quantifiers purports to provide a sharp diagnosis of the invalidity of the ontological argument for the existence of God, along the same lines but more precisely than that advanced by Kant. The argument tells us that the statement ‘God exists’ is true by the very meaning of the word ‘God’ and cannot be denied without contradiction. In particular, ‘God’ is defined as that being that satisfies the following predicates (presumably there are more predicates that serve to characterise God, but we’ll just stick to these three):

a is omniscient.

a is omnipotent.

a exists.

So it is actually contradictory to deny the statement that God exists; to do so is like denying that a square has four sides. If we insist that existence is actually to be expressed as a quantifier and not as a ‘first-level’ predicate, however, the last member of the list of predicates is ill formed. The statement that God exists must be written:

(17) There is an x such that (x is omnipotent & x is omniscient).

One who denies this does not thereby get into logical trouble.

We’ll come back to this in the chapter on Russell.²

● SUBSTITUTIVITY AND EXTENSIONALITY

As explained above in the section ‘Compositionality again; the reference of a sentence’, Frege splits the principle of compositionality into two. The principle of compositionality of reference tells us that if we begin with a sentence S and replace a part of S with an expression with the same reference, then the resulting sentence will have the same truth-value as S . We can state this more accurately:

The principle of substitutivity: If S contains expression e , and e^* has the same referent as e , and S^* results from S by replacing e by e^* , then S^* has the same referent as S .

If S and S^* are sentences, each of whose referents is a truth-value, then substituting co-referring parts does not change the sentence’s truth-value (replacement of co-referentials saves truth, or *salva veritate*). It is important to appreciate how deeply founded this principle is in our understanding of language. Consider the following:

(18) Hesperus is a planet.

(19) Venus = Hesperus.

(20) Venus is a planet.

The truth of (19) shows us that ‘Venus’ and ‘Hesperus’ are co-referential: Venus and Hesperus are the same object. Therefore (18) and (20) say the very same thing about the very same object. So they *cannot* differ in truth-value. What they say about the object is what is expressed by ‘ α is a planet’, and the object they say this about is Venus, otherwise known as Hesperus. The principle of substitutivity seems to get to the root of the very idea of what it is to talk *about* an object, of *reference* itself. If one’s aim is to say something true about an object, it doesn’t matter by what means we refer to the object; it doesn’t matter how we *pick out* the object. All that matters is whether that object is as we say it is; if it is, then what we say is true, and if not, then what we say is false.

The principle of substitutivity is not restricted to any particular kind of expression. Consider predicates. Suppose that all and only the members of a certain club slept in the mansion on the night of the murder. Then if Smith is a member of the club, it follows that he slept in the mansion on the night of the murder; similarly, if Jones did not sleep in the mansion that night, then he is not a member of the club. Put in terms of the principle of substitutivity, this is to say that since ‘ α is a member of the club’ and ‘ α slept in the mansion on the night of the murder’ are co-extensive, Smith’s satisfying one predicate entails that he satisfies the other, and Jones’ not satisfying one entails that he doesn’t satisfy the other.

Co-extensive predicates are interchangeable in other contexts, too, contexts in which the predicate occurs in a non-atomic sentence. For example, the following predicates are co-extensive (assuming that so-called flying squirrels do not really *fly*; they only *glide*!):

α is a bat.

α is a flying mammal.

Their co-extensiveness entails the truth of:

For every x (x is a bat iff x is a flying mammal).

That is, something is a bat, if, and only if, it is a flying mammal.

A language in which the principle of substitutivity holds generally is called an **extensional** language. A language in which the principle fails at least some of the time is a **non-extensional** language.

• THE ANALYSIS OF PROPOSITIONAL ATTITUDES

The principle of substitutivity seems utterly axiomatic. Nevertheless, it might seem that there are counterexamples. Consider the following pair:

(21) Sally believes that Venus is a planet.

(22) Sally believes that Hesperus is a planet.

Clearly, it would be possible for (21) to be true but (22) false, if Sally did not happen to know that Venus is Hesperus. (In fact, Sally could positively believe that Hesperus is *not* a planet yet believe that Venus is a planet, without being irrational or illogical in any way.) Yet ‘Hesperus’ and ‘Venus’ are co-referential; so how could (21) and (22) differ in truth-value? Both (21) and (22), one might think, say the same thing about Venus: they both say of that object that it is *believed by Sally to be a planet*.

The answer in terms of Frege’s theory is that (21) and (22) are not really *about* Venus at all (though they are about Sally). The predicate ‘Sally believes that α is a planet’ is not about planets. The truth-value of a sentence formed from it does not depend on what if anything is referred to by the singular term slotted into it. This is plausible because (22) could be true even if ‘Hesperus’ did not refer. Pretend that actually there is no such planet. If the referent-determining rule expressed by ‘Hesperus’ did not happen to pick out an object, then still ‘Hesperus is a planet’ would express a proposition, and Sally might believe that proposition. As we saw, that is exactly what happened in the case of the French astronomer Le Verrier and the supposed planet Vulcan in the nineteenth century.

(21), then, is not about Sally and Venus, but about Sally and the proposition that Venus is a planet – similarly for (22), Sally and Hesperus. This sort of fact is reflected in the difference between:

(23) Venus is a planet.

(24) that Venus is a planet.

Consider the proposition expressed by (23). (23) *expresses* that proposition, but (24) *refers* to that proposition. (24) *is a singular term that refers to the proposition expressed by* (23). For Frege, in general:

Principle of that-clauses: the result of prefixing a declarative sentence S by ‘that’ is a singular term whose referent is the sense of S .

What (21) says, then, is that Sally stands in a certain relation to a certain proposition (Frege’s term is a certain *thought*). In particular, she stands in the *believing* relation to the proposition that Venus is a planet. At the first level of analysis, the sentence decomposes into a singular term standing for Sally, a two-place predicate ‘ α believes β ’ and a singular term standing for the proposition that Venus is a planet. It is of the same form as ‘Sally kissed John’. This is exactly what we should expect, given our earlier characterisation of the propositional attitudes: belief is a certain relation between a believer and a proposition. All of this carries over to the other propositional attitudes (fear, hope etc.).

One way to think about this, then, is that the possibility of (21)’s being true but (22)’s being false is a counterexample to the principle of substitutivity.³ A language that has such devices as the sentence-operator ‘that ____’ is not therefore extensional. It is ‘non-extensional’: the referent of a that-clause does not depend on the *referent* of the sentence slotted in – a truth-value – but on the *sense* of the sentence slotted in. The role of ‘Venus’ and ‘Hesperus’ in (21) and (22) is not to refer to Venus; as philosophers sometimes say, that-clauses create a non-extensional *context*.

The referent of (23), ‘Venus is a planet’, is a truth-value, but the referent of (24), ‘that Venus is a planet’, is the sense of ‘Venus is a planet’. The sense of the sentence becomes the referent of the corresponding that-clause; Frege also calls the sense of the sentence its *indirect referent*. What is the *sense* of ‘that Venus is a planet’? It cannot be the same as the sense of ‘Venus is a planet’ – for sense determines reference, and these have different referents. Frege calls the sense of ‘that Venus is a planet’ the *indirect sense* of ‘Venus is a planet’. It is a higher-order sense, or level-two sense. And this accounts for the difference of *sense* between (21) and (22), and thus accounts for their possible divergence in *reference*, their possible divergence in truth-value.

We’ll return this topic in Chapter 11.

● THE OBJECTIVITY OF SENSE

A sense of a referring term is both a *rule for determining* the reference of the term and a *mode of presentation* of the reference (if it has one). In an especially lucid passage in ‘On Sense and Reference’, Frege writes:

The following analogy will perhaps clarify these relationships. Somebody observes the Moon through a telescope. I compare the Moon itself to the reference; it is the object of the observation, mediated by the real image projected by the object glass in the interior of the telescope, and the retinal image of the observer. The former I compare to the sense, the latter is like the idea or experience. The optical image in the telescope is indeed one-sided and dependent upon the standpoint of observation; but it is still objective, inasmuch as it can be used by several observers... But each one would have his own retinal image.

(‘On *Sinn* and *Bedeutung*’, Frege 1997, p. 155; translation slightly altered)

In ‘The Thought: A Logical Inquiry’, Frege again stresses that sense – what he calls the thought, what we are calling the proposition – is not mental, not experiential, not psychological, not of the same ilk as ideas. A proposition, as we said, is the kind of thing that can properly be said to be true or false. It is the objective meaning of a sentence and is the object or content of a propositional attitude (belief, wish etc.). According to Frege, it follows from these considerations that propositions cannot be mental or psychological entities.

For what is a mental entity, an *idea* in the sense in which Locke used that term? One example would be the sort of *mental image* we have when imagining something or when we dream. Such a thing, as many philosophers have pointed out, is *private*. Suppose I am imagining my grandmother (now deceased). I have a vivid mental image in my mind’s eye. It seems possible to describe to you what this image is like in such a way that you can generate your own image that is perhaps like it, but *you cannot perceive my mental image*. You cannot, in principle, verify that your image is like mine. Only I can perceive or experience my mental image: it is private. Likewise, only you can perceive yours. Mental images, and mental entities more generally, are

mental in the sense that they are in the mind: the cash value of this seems to be that they are *private*, necessarily available only to the mind they are in.

Consider now the hypothesis that the meaning of a sentence is a mental entity – an idea, as Locke put it, or a mental image, as we might more elaborately put it. I say: ‘Bologna is north of Florence’. What I mean is private to me, shut up in my mind. How can you understand me? How can you grasp *what I mean*? Well, you hear the words, which cause, in your mind, some kind of mental image. But how are you to know that this mental image has *anything at all* to do with mine? What could tell you that it is similar to mine? You might try to describe your mental image, asking me whether yours is like mine. But these are just more words: you attach a mental image to them, and they cause a mental image in me. How are we to find out whether *these* mental images are at all alike? Clearly, there is no way: we can repeat the game forever and never get any closer to finding out whether your thoughts are anything like mine. As Frege put it, language on this model would not amount to mutual understanding at all; utterances would at best be caused by mental states, not properly expressing a meaning at all.

If there is to be communication at all, meaning must be something ‘out there’ in some way: instead of being private, it must be something public, something that is no more my possession than yours. Of course, there are mental states; there is such a thing as the psychological process of thinking; there are mental processes that go on when we mean something by what we say. Nevertheless, a sentence like ‘Bologna is north of Florence’ has a public meaning: when uttered in the normal way, there are linguistic rules that determine it as meaning that a certain city stands in a certain geographical relation to another. The individual person, when hearing it spoken, may understand it – recognise the proposition it expresses according to compositional rules – or fail to understand it. This holds even when we are talking about inner states. If Sally says ‘I am hungry’, then according to linguistic rules, this sentence is true if and only if the speaker is in a certain physiological state at the time of utterance. If she says ‘I am sad’, then it is true if and only if the speaker is in a certain mental state at the time of utterance. The sentence is *about* a mental state, but its *sense* is not a mental state.

Fregean sense, then, is not mental or psychological. Of course, it isn’t physical either: senses such as propositions are not physical objects. They are reference-determining rules associated with expressions. Sense, then, inhabits the strange world of the *abstract*: senses, like mathematical objects, are non-spatial objects with no causal powers.

Finally, it is important to be clear that *sense*, in Frege’s terminology, *has nothing to do with the senses*, such as seeing and hearing. Also, it has nothing to do with *sense-data*, which Russell and others regarded as the immediate objects of perception. There is absolutely no presumption that a word with sense has to appertain to something *sensible*, something that can be sensed or perceived. For example, ‘ $2 + 2 = 4$ ’ expresses a sense, but the sense has nothing to do with anything sensible.

This should strike one as perfectly natural: we do ordinarily speak of the *sense* of a sentence and describe meaningless talk as *nonsense*.

● PREDICATE REFERENCE AND THE CONCEPT HORSE PROBLEM

Frege's view of the *senses* of predicates is intuitive: for example, to understand the sense of the predicate ' α is a horse' is to understand the rule or criterion for whether or not the predicate applies to an arbitrary thing. If one does grasp that sense, and also grasps the sense of the singular term 'Red Rum', then one grasps the proposition expressed by 'Red Rum is a horse'.

The *reference* of predicates is another matter. Earlier, I said that for the sake of simplicity we would suppose that Frege thought of the referents of predicates as their extensions, the set of things that the predicate applies to. Frege's real view is more complicated (the classic statement is in 'On Concept and Object').

Consider:

(25) Red Rum is a horse.

We can't really take the referent of ' α is a horse' to be the set of horses, for that set is merely an *object*, nameable by a singular term – i.e. 'the set of horses'. What is so special about those particular objects which makes them suitable for references of predicates, for making propositions? Why those and not, say, numbers, or pencils? If one says 'Red Rum, my pencil!', one lists things; one doesn't express a proposition – similarly with 'Red Rum, the set of horses'! If we try saying (25) is really

(26) Red Rum *is a member of* the set of horses.

we thereby launch an infinite regress: that sentence mentions Red Rum and the set of horses, but it *also* includes the words 'is a member of'. If we explained that expression, like the original ' α is a horse', as itself referring to a set, the same issue would arise. We'd get something like the monstrous

(27) Red Rum the set of ordered pairs such that the first is a member of the second the set of horses.

This, again, is a list, not a sentence. *The moral is that a sentence must have something whose role is not merely to refer to an object.*

Instead, Frege holds that a predicate denotes a *function*, a kind of entity whose mode of being is quite different from that of objects. Just as the value of the function denoted by ' $\alpha + 2$ ' when applied to 3 is 5, and the value of the denotation of 'the mother of α ' for a person is the person's mother, the value of the function denoted by ' α is a horse' for Red Rum is either of the two truth-values, depending on whether or not Red Rum is a horse. Since, in fact, Red Rum is (or was) a horse, the truth-value of the sentence is truth.

Frege calls functions whose values are truth-values **concepts**. The terminology is unfortunate because it seems more appropriate to call the *senses* of predicates by

that word, as, indeed, did the great Fregean Alonzo Church. But in the remainder of this chapter we'll stick with Frege's way of speaking.

Fregean concepts and functions in general are not objects and cannot be referred to by singular terms. They are 'second-level' entities. Objects comprise the *first-level*, and entities that can be applied to objects are *second-level* entities.⁴ Ignoring functions like that denoted by ' $x + 2$ ', which are not concepts, we can say equally that first-level entities, objects, can be the inputs for second-level entities, the concepts, which yield truth-values as outputs. Functions are like arrows, pointing from one entity to another; concepts are the special case where the entities pointed to are the truth-values, truth or falsity. Predicates refer to them, such as:

α is green.

α is a friend of the Dalai Lama.

$\alpha + 2 = 7$.

But Frege now perceives a riddle. The referent of a predicate such as

(28) α is a horse

is a concept. Now consider:

(29) ' α is a horse' refers to the concept horse.

(30) 'the concept horse' refers to the concept horse.

(29) seems unobjectionable. And (30) seems inevitable – surely instances of the schema " A refers to A " will be true no matter what referring term we put for ' A '. Thus the two terms ' α is a horse' and 'the concept horse' should be thought of as referring to the same entity. Therefore, at least in extensional contexts, they should be intersubstitutable. Wherever one occurs (extensionally), it should be possible to replace it with the other without changing the truth-value of the whole in which it occurs. Continue with an empirical truth:

(31) Red Rum is a horse.

Thus, from the point just made that ' α is a horse' and 'the concept horse' being inter-substitutable, we infer:

(32) Red Rum the concept horse.

But again, this is a list, not a sentence. Yet it gives the appearance of following from true premises.

The problem, in Frege's view, is with the phrase 'the concept horse'. (29) is *not* true. ' α is a horse' has a gap in it indicated by the Greek letter, whereas 'the concept

horse' does not. It isn't the sort of thing that can play the role of a predicate; predicates are necessarily *incomplete*, or *unsaturated*, as Frege puts it, and concepts by their nature are incapable of being referred to by terms that are complete, saturated, or gapless.

What should we write in place of (29)? Nothing. For Frege, *predicates only refer when they are performing their normal roles in complete sentences*. Indeed, it seems that any attempt to say what predicates refer to is doomed to failure – since grammar itself seems to require that the space in 'A refers to ___' is replaceable only by a singular term, which thus refers to an object, a first-level entity, not a concept, not a second-level entity. It is an extremely paradoxical situation, which has inspired many, many responses. Frege himself was well aware of the situation and begged the reader to grant him a 'grain of salt'.

● FURTHER DISCUSSION: THE CONTEXT PRINCIPLE

In a famous passage of the *Foundations of Arithmetic*, Frege enunciates three methodological principles which he strives to obey. The second has come to be known as the context principle (where a 'context' means a location in a sentence, not a context of utterance including time, place and the identity of speaker):

[one should] never ... ask for the meaning of a word in isolation, but only in the context of a proposition.

(Frege 1974 [1884], introduction)

This was before he had drawn the distinction between sense and reference, but – although it has been the subject of some controversy – it remains plausible that we can bifurcate the principle as (i) one must not ask after the reference of a word in isolation but only in the context of the truth-values of whole sentences, and (ii) one must not ask after the sense of a word in isolation but only in the context of the senses of whole sentences. The senses of whole sentences in Frege's terminology are *thoughts*, which according to our terminology are *propositions* (his use of 'proposition' in the above statement is roughly equivalent to 'significant sentence').

Assume that to specify a proposition is to specify a truth-condition (so the truth-condition of 'Snow is white' is that snow is white). The principle seems to tell us that the sense of a singular term (for example) cannot be properly conceived apart from the effect it has on the truth-conditions of whole sentences. Indeed, if the senses of all sentences in which the singular term occurs are settled, then no further question remains as to the sense of the singular term. Since the sense of a singular term determines the term's reference – if it has one – nothing further from the point of view of language remains to be done for settling the reference of a singular term. All that remains is that the extra-linguistic facts should cooperate, to avoid the fate of Le Verrier and his belief in the existence in the planet Vulcan.

It might well seem that if you have to teach someone a word, you do it by teaching the person to use the word in whole sentences, perhaps by giving the person examples of whole sentences involving the word. Even if the method involves pointing to an object and saying the word, perhaps what is being conveyed is really a sentence of the form ‘That [pointing] = A’, where ‘A’ is the singular term being introduced. And it makes perfect sense of reference to things that *cannot* be pointed to – i.e. abstract objects (such as numbers, which are most relevant to a book on the foundations of arithmetic). Also, (ii) – along with (i) – seems to make philosophical sense of the idea of the section ‘Predicate reference and the concept horse problem’, above, that ‘predicates only refer when they are performing their normal roles in complete sentences’. A reason is that predicates have sense only insofar as they contribute to the senses of whole sentences (some scholars think that the point holds also for singular terms).

But (ii) is somewhat unsettling in its implications. It means that linguistic reference cannot exist without mastery of a substantial chunk of language. A child, for example, who has only a few words, including ‘Fido’ for the family dog, cannot strictly speaking be credited with *referring* to Fido by their use of ‘Fido’ (though perhaps they can be characterised as a ‘proto-referrer’). But how then does the child acquire language to begin with? If they cannot learn words without having already mastered sentences, and sentences, of course, are made up of words, then the learning of language looks impossible. In particular, it looks as if there is a conflict between two Fregean principles:

Compositionality: the senses of whole sentences are determined by the senses of the words and the way they are put together.

Context: the senses of words are determined by the senses of whole sentences in which they occur.

But perhaps the conflict is unreal. Perhaps, at the early stages, language is not learned one word at a time, as if one had to learn the complete sense of word before moving to the next, but a substantial portion of language gradually and simultaneously comes into focus as whole. Progression is not direct, a matter of learning words one at time, but crab-like, or like a person climbing up with grappling hooks.

● HISTORICAL NOTES

Frege was not well known until long after he died in 1925, although philosophers Bertrand Russell, Rudolf Carnap and Ludwig Wittgenstein were all keenly aware of him while he was alive, and Edmund Husserl had some contact with him; and among leading pure mathematicians, Giuseppe Peano and David Hilbert corresponded with him. Most people say he was without a doubt the unsung inventor or discoverer of modern logic, the symbolic logic which handles inferences involving relations (a simple example is ‘Every mollusc is an animal; therefore all who eat molluscs are eaters of animals’). The key – which is now standard in logic – is the integration of sentential logic with variables, predicates and quantifiers, as described briefly in

the last chapter. This advance was modestly but firmly announced in *Begriffsschrift*, which appeared in 1879. The logic was vastly more powerful than anything that existed before, and it made the thesis plausible that arithmetic is really logic – a thesis that he sketched informally in the *Foundations of Arithmetic* of 1884 and attempted to prove rigorously in the *Basic Laws of Arithmetic* of 1893 and 1903. The proofs and presentation were impeccable, except for one detail: his assumption that every predicate determines an extension, or class, or set, was inconsistent. Famously, Russell wrote to Frege informing him of this fact, just as Volume 2 of the *Basic Laws* was to be printed. Frege arguably never recovered from what he described as the ‘shock’ of having received the news from Russell; it was for Russell, along with A. N. Whitehead, to make the case in a more complicated way for the reducibility of arithmetic to logic, in *Principia Mathematica* of 1911. See Chapter 4 for more. Husserl’s ideas on perception as expressed in his great *Logical Investigations* of 1900 (second edition 1913) may have been influenced by his correspondence with Frege (see question number 6 below); more palpable though controversial was the influence on his view of arithmetic.

The exact connection of Frege’s work in logic and the foundations of mathematics to his ideas in the philosophy of language has been the subject of some dispute. But among the undisputed elements is that an explicit distinction between sense and reference was announced in 1892, with ‘On Sense and Reference’. It enabled him to maintain rigorously that pure mathematics delivers genuine knowledge, despite its consisting of logically true equations. No less important in Frege’s eyes was ‘On Concept and Object’, also from 1892; the central idea of that piece had been active in his thought since 1884 at the latest, but the power and precision of his analysis on such an abstract topic was almost without precedent. After 1903, little was heard from Frege until 1918, when he published ‘The Thought: A Logical Inquiry’, in which he makes the case for the ‘third realm’, the objectivity and abstract nature of propositional content. Russell thought Frege’s sense-reference distinction unnecessary. In a way, Wittgenstein agreed with Russell on that point, but there is no mistaking Wittgenstein’s debt to Frege in his early work the *Tractatus Logico-Philosophicus* (Wittgenstein says that Frege ‘quite wiped the floor’ with him during his 1912 visit to Frege, on the recommendation of Russell). Carnap developed Fregean themes in his *Introduction to Semantics* (1942) and *Meaning and Necessity* (1956 [1947]); he had attended courses with Frege in the early 1910s. Michael Dummett (1925–2011) did the most to establish Frege’s name at the top of analytic philosophy, with a series of great books about him starting with *Frege: Philosophy of Language* (1973).

● CHAPTER SUMMARY

Frege begins with a two-pronged attack on naïve semantics, focusing on (NP2): the meaning of a singular term is its referent.

- 1 The first prong concerns co-referential singular terms. For example, since the Morning Star = the Evening Star, the meaning of ‘the Morning Star’ = the

meaning of 'the Evening Star', according to the naïve theory. But that conflicts with the evident fact that the two sentences 'The Morning Star is a planet' and 'The Evening Star is a planet' do *not* mean the same. The two sentences do not have the same cognitive value, as Frege puts it; discovering that one is true is not the same as discovering the truth of the other. It also conflicts more dramatically with the observation that whereas 'the Morning Star = the Morning Star' is *a priori*, 'the Morning Star = the Evening Star' is empirical, *a posteriori*.

- 2 The second prong of the attack concerns non-referring singular terms, such as 'Vulcan' or 'The Loch Ness Monster'. According to naïve theory, their lacking reference entails that they lack meaning; but this flies in the face of the evident difference in meaning between 'Vulcan is hot' and 'Nessie is hot', as well as the fact that a person can believe the content of 'Vulcan is hot' while disbelieving that of 'Nessie is hot'.

Frege responds to this by postulating that each meaningful expression expresses a sense, which is a rule for determining the referent of the expression, if it has one. If the sense does determine a referent, then the sense acts as a mode of presentation of the referent. The distinction is drawn across the board: a singular term expresses a sense that purports to determine an object as its referent; a one-place predicate expresses a sense that purports to determine an extension (in fact a *function*) as its referent, and similarly for two-place predicates and so on; and a sentence expresses a sense (a proposition, what Frege calls a thought) that purports to determine a truth-value as its referent. Thus the principle of compositionality is split in two, one for reference and one for sense (there is also a third, operating on syntax).

Thus (1) 'The Morning Star is a planet' and 'The Evening Star is a planet' have different senses despite their saying the same thing about the same object. And 'Sally believes that the Morning Star is a planet' and 'Sally believes that the Evening Star is a planet' need not coincide in truth-value, for the two relate Sally to different propositions. (2) 'Vulcan is a planet' is neither-true-nor-false, because of the non-existence of Vulcan; but it expresses a proposition despite lacking a truth-value. Therefore 'Sally believes that Vulcan is hot' has a truth-value. However, Frege still has the problem of negative singular existentials: there appears to be no non-arbitrary way of explaining how 'Vulcan does not exist' is true; Frege's view gets the wrong answer that it is neither-true-nor-false.

Frege argues that there is no coherent alternative to the objectivity of sense; propositions are not in anyone's mind but are public abstract entities.

Frege takes predicates to refer to unsaturated entities he calls 'concepts'. Any sentence – take 'Jane smokes' as a representative – must have among the entities to which it refers an unsaturated or incomplete entity in addition to objects like Jane, for taking the rest of the sentence to denote another saturated or complete entity – an object – would have it as only a list of objects, not a sentence. Therefore the sentence must be taken to have the referent Jane and the unsaturated referent of ' α smokes', a Fregean concept. This latter entity is for Frege a function: it has as its value the truth-value truth if Jane smokes, and falsehood if Jane does not smoke. The

Fregean puzzle of the concept horse arises because of the apparent reasoning from a true statement to something that does not even express a proposition: Red Rum is a horse (suppose true); ‘ α is a horse’ refers to the concept horse; therefore Red Rum the concept horse. Frege diagnoses the problem as that “‘ α is a horse’ refers to the concept horse” is, strangely enough, not true; ‘the concept horse’, in being a saturated expression, refers at most to an object, not to an unsaturated entity, a concept.

Frege espoused the so-called context principle, which was ‘never to ask for the meaning of a word in isolation, but only in the context of a proposition’. Exact interpretation of the principle has been somewhat contentious, but it is undeniable that Frege appealed to it in support of his thesis that we refer to abstract objects (including numbers). The principle is sometimes held to conflict with another principle of his, the principle of compositionality: on the face of it, context says sentence-meaning is prior to word-meaning; composition says word-meaning is prior to sentence-meaning.

● STUDY QUESTIONS

- 1 Assume that Johnny sincerely exclaims: ‘Santa Claus is in the department store!’ Suppose that there is a man dressed up in a red suit etc. in the department store. Is what Johnny says false? True? Does Johnny falsely believe that Santa Claus is in the department store? What is the meaning of ‘Santa Claus’ in terms of Fregean semantics?
- 2 What do you think of the following strategy for avoiding the problem of negative singular existentials within Fregean semantics? Introduce the expression ‘It’s true that P’, which says of a *proposition* P that it’s true; if P is false *or defective in such a way as to be neither true nor false*, it yields a falsehood; and if P is true, it yields a truth. Then analyse ‘Vulcan does not exist’ as ‘Not: It’s true that (there is an x such that x = Vulcan)’.
- 3 It is plausible to say that ‘bought’ and ‘purchased’ are synonyms – that they express the same sense. Is the following therefore a valid argument? Why or why not?

Sam bought a turnip.

Therefore, Sam purchased a turnip.

What about:

Susie believes that Sam bought a turnip.

Therefore, Susie believes that Sam purchased a turnip.

Discuss the relation between:

that Sam bought a turnip

the sense of ‘Sam bought a turnip’

- 4 *Conceptual analysis* is, in Fregean terms, the search for statements like:
- (a) *x knows that p if and only if x has a justified true belief that p*,
where the parts italicised express the very same sense. If it is successful, then apparently such a statement is the same statement as
 - (b) *x knows that p if and only if x knows that p*.
But (b) is a triviality, which cannot extend our knowledge; (a), even if it does represent a successful analysis, seems not to be a triviality. This is known as the *paradox of analysis*; what should a Fregean say in response?
- 5 For Frege, except in cases where the expression fails to determine a referent, the sense of an expression is a mode of presentation of its referent. If we accept this, do we say that psychological acts of thinking necessarily involve senses and not necessarily referents? Compare these ideas with the idea that the content of a perception is never simply an object but either (a) an object *from-a-particular-point-of-view* or (b) a collection of sense-data, which could be exactly the same if it were a case of hallucination and the object did not exist. (This is one connection with Husserl, who spoke of the 'noema' of an act of perception, which is in some way like Fregean sense.)
- 6 For Frege, concepts are a subclass of functions (remember that in Fregean language, concepts are the referents of predicates, not their senses). Therefore just as the function ' $\alpha + 5$ ' has the value 7 for the argument 2 (i.e. ' $2 + 5$ ' is a singular term denoting 7), ' α is white' has the value true for the argument snow (i.e. 'Snow is white' is a singular term denoting truth). Are all true sentences thus singular terms denoting truth-values? Are truth-values objects, alongside the Eiffel Tower? Is there no logical difference between sentences and singular terms? Does that mean that one can assert something by saying 'Truth!?' or 'The Eiffel Tower!?' Can you think of reasons for doubting this?

● PRIMARY READING

All relevant sources for Frege are included in *The Frege Reader*, edited by Michael Beaney (1997); the main ones are 'On *Sinn* and *Bedeutung*' ['On Sense and Reference' 1892], 'On Concept and Object' [1892] and 'Thought' [1918]; see also 'Function and Concept' [1891], the 'Letter to Husserl, 24.5.1891', the '[Comments on *Sinn* and *Bedeutung*]' [1892], the '*Grundgesetze der Arithmetik*, Volume 1 (1893): Selections', the 'Letter to Russell, 13.11.1904' and the excellent 'Letter to Jourdain, Jan. 1914'. Beaney has cleverly finessed the delicate matter of whether to translate *Bedeutung* as 'reference' or 'meaning' by not translating it. Although it was written well before he had the idea of distinguishing sense from reference, no set of recommendations for reading Frege would be complete without mentioning Frege's wonderful *Foundations of Arithmetic*.

● SECONDARY READING

There are now many books about Frege, but certainly the most substantive is M. Dummett, *Frege: Philosophy of Language*, second edition (1993). The book is

outstanding, but you may want to begin with two essays by Dummett: ‘Frege’s Philosophy’ and ‘Frege’s Distinction Between Sense and Reference’, in Dummett’s collection *Truth and Other Enigmas* (1978).

A much shorter and more accessible book is H. Noonan, *Frege: A Critical Introduction* (2001).

• NOTES

- 1 Experienced Fregeans will recognise a number of places in which I simplify or omit. Some Fregeans doubt that his theory was meant to apply to *natural language*, that his remarks were only to explain his artificial symbolic language, his *Begriffsschrift* (concept-script). I think he thought that his *Begriffsschrift* more exactly mirrors what goes on more vaguely in natural language, namely the expression of thoughts, of propositions.
- 2 Frege himself, in his great work the *Basic Laws of Arithmetic*, accommodated this shortcoming by stipulating that in such cases, in effect, the referent is actually the empty or null set. But as Frege himself allows, this is plainly artificial. At any rate, we don’t want to say that Vulcan is the empty set, or that Vulcan = the present King of France.
- 3 An alternative interpretation is that Frege did believe that the principle of substitutivity holds everywhere, and that extensionalism holds everywhere. Reference, on this view, is relative to context: in ordinary contexts like (21), the referent is the direct referent. In contexts like (22), the referent is the indirect referent. Indeed, this is probably closer to Frege’s thinking, and it fits with his context principle (§10). But I think it clearer to the student to characterise Frege as holding that reference is not relative in this way.
- 4 There are also third-level entities and fourth-level entities and so on. Quantifiers, since they take second-level entities (concepts) as arguments, are third-level. There are even mixed-level entities, such as the referent of ‘ α falls under the concept- ϕ ’.

3

• russellian semantics

• THE TASK FOR RUSSELL

Two sorts of problem motivated Frege's distinction between sense and reference: problems of empty singular terms and problems of cognitive value. For the purpose of discussing Russell, we shall rename them as the problems of:

- 1 existence
- 2 identity

Existence problems include the question of how to explain the evident meaningfulness of sentences like:

The man who killed the Kennedys worked for the Mafia.

Vulcan is hot.¹

– despite the fact that there is no such man as the man who killed the Kennedys (no single person killed both Bobby and John) and no such planet as Vulcan, in which case the singular terms 'The man who killed the Kennedys' and 'Vulcan' are empty – i.e. do not refer to anything. Note that such a sentence as this could be true:

Jones believes that Vulcan is hot.

– which implies that there is such a proposition as that Vulcan is hot, in which case 'Vulcan' must be meaningful, despite not having a referent.

There is a crucial, further variety of existence problem to which Frege did not explicitly offer a satisfactory response (though see Question 2, p. 66). For example:

The Loch Ness Monster does not exist.

Vulcan does not exist.

are both not only meaningful but *true*. But then the semantical instructions cannot plausibly be: find the referent of the singular term – ‘The Loch Ness Monster’ or ‘Vulcan’ – and determine whether or not it satisfies the predicate ‘does not exist’ (just as they tell us, in the case of ‘Seattle is rainy’, to find the referent of the singular term – ‘Seattle’ – and find out whether it satisfies the predicate – ‘is rainy’). For in the case of ‘The Loch Ness Monster’ or ‘Vulcan’, there is no referent, therefore nothing to check for satisfaction or non-satisfaction of the predicate.

Identity problems concern the fact that co-referring singular terms may differently affect what Frege called the *cognitive values* of sentences in which they occur. The most glaring examples include pairs of such truths as:

Venus = Venus.

Venus = the Evening Star.

This is the most glaring kind of case because the first sentence is so unlike the second in cognitive value. A case like the following shows that sentences can say the same thing about the same object but differ in cognitive value:

Venus is a planet.

The Evening Star is a planet.

Since Venus is identical with the Evening Star, a theory that tries to explain meaning in terms of reference alone, such as the naïve theory, seems quite unsatisfactory.

Because of this, sentences that ascribe propositional attitudes may present apparent failures of the **substitutivity** principle, as in:

George wondered whether Scott = the author of *Waverley*.

George wondered whether Scott = Scott.

According to the example from Russell, the Prince Regent George, later King George IV, did wonder whether it was the poet Walter Scott who wrote the famous Scottish novels (in fact he did); but he did not of course wonder whether Scott is Scott.

Similarly, the following sentences need not have the same truth-value:

George believes that Venus is a planet.

George believes that the Evening Star is a planet.

But Russell does *not* accept Frege’s distinction between sense and reference. He thinks there is a more economical way to solve the existence and identity problems.

Russell accepts the main principle of naïve semantics, that, essentially, meaning is just reference, but he finds ways around the apparent need to postulate sense.

● THE THEORY OF DEFINITE DESCRIPTIONS

We can distinguish between *simple* singular terms and *complex* singular terms. Simple singular terms are proper names, such as 'Paris', 'Michelangelo' and so on (this includes names such as 'Winston Churchill', which comprise more than one word; we shall ignore this complication and regard all proper names as simple singular terms). Complex singular terms include such examples as:

the capital of France

Alexander's father

the most intelligent human being

An important fact is that *all* complex singular terms can be rewritten in the form

the F

where F is some predicate. For example, the singular term 'Alexander's father' can be written as 'The father of Alexander'.²

A singular term of the form 'the F' is what Russell calls a **definite description**. His theory of descriptions is an account of the word 'the'. According to what we said in Chapter 1, a predicate is what remains when we remove a singular term from a sentence. Thus any sentence containing a definite description 'The F' can be understood as having the following form (for some possibly complex predicate ' α is G'):³

(1) The F is G.

Thus consider an example:

(2) The present King of France is wise.

According to Frege's theory, this sentence is meaningful – that is, it expresses a sense – but it is neither true nor false, since 'the present King of France' does not have a referent. But Russell thinks that that cannot be right, for a meaningful sentence, a sentence expressing a sense, as Frege would put it, expresses a *truth-condition*, a way that the world must be if the proposition is to be true. Either the world is that way, in which case the sentence is true, or it is not, in which case the sentence is false. (2) *does* express a truth-condition: therefore, since it is not true, it must be false. For what could it mean to say that a sentence is false other than that it is meaningful (expresses a proposition) but is not true?

The truth-condition of (2) is clear. It may be expressed as a conjunction of three propositions – the *existence*, *uniqueness* and *categorical* clauses:

- (2a) There is a King of France.
- (2b) There is not more than one King of France.
- (2c) Any King of France is wise.

A little reflection shows that (2a)–(2c) are jointly sufficient and individually necessary for the truth of (2). Their conjunction is necessarily equivalent to (2). We can put the point more generally by saying that (1) is necessarily equivalent to the conjunction of:

- (1a) There is an x such that Fx .
- (1b) Not: (there is an x and there is a y such that ($x \neq y$ & Fx & Fy)).
- (1c) For every x ($Fx \rightarrow Gx$).

(1) can be understood as saying ‘There is exactly one F , and every F is G ’. In fact, the conjunction (1a–1c) is logically equivalent to the more compact:

- (1*) There is an x such that (Fx & for every y ($Fy \rightarrow y = x$) & Gx).⁴

Russell’s claim is that sentences of the form (1) are analysed, say the same thing as, sentences of the form (1*).

Russell stresses the affinity between ‘The F is G ’ and ‘There is an F that is G ’ (‘Some F is G ’): whereas the former, in Russell’s view, says that there is exactly one F and it is G , the latter says that *at least* one F is G . In other words, ‘The’ should be considered a quantifier, along with ‘There is’ or ‘Some’ – and also alongside ‘All’ and ‘Every’ and so on (called ‘determiners’ in the parlance of linguistics). If we consider the phrases

The F
 Some F
 An F
 All F
 Each F

and so on, it is clear that any one of them forms a sentence when joined to a predicate ‘... is G ’ (or ‘... are G ’), and that each has a meaning irrespective of whether the predicate F is satisfied by any objects. A way to put it is that they are all *descriptions*, but the first is a *definite* description and the rest are *indefinite* descriptions.

This observation helps to dislodge the naïve idea that ‘The F ’ is a genuine singular term – i.e. a referring expression like the name ‘Sigmund Freud’ – despite the fact that both can slot into the empty place in a one-place predicate to form a sentence.

'The F is G', as shown by its analysis as (1*) or the conjunction (1a)–(1c), has nothing like the simple form 'Fa' ('a is F'). Unlike (1*) or the conjunction (1a)–(1c), 'Fa' is the simplest possible sentence, an atomic predication. The *logical form* of 'The F is G' is very different from its *surface form* – i.e. its grammatical or linguistic form.

In accord with this, Russell's way of explaining the word 'The' is what he calls **contextual definition**, or *definition-in-use*. An **explicit** or **direct** definition substitutes a given symbol with another that is grammatically of the same category (the defined expression is a simple symbol, and the defining expression normally is complex). For example, we can define 'bachelor' as 'unmarried man'; these are equivalent expressions of the same grammatical category:

df: For every x, x is a bachelor if and only if x is an unmarried man.

Wherever the symbol 'unmarried man' occurs, the definition licenses one to replace the symbol with 'bachelor' (or vice versa).

By contrast, the technique of contextual definition of a symbol gives a rule for transforming *whole sentences* containing that symbol into whole sentences that do not contain the symbol. So Russell does not give us a form of words or symbols that is equivalent to 'The' or to 'The F'. Rather, he explains how to re-express the content of any sentence of the form:

The F is G.

without substituting some other kind of singular term for 'The F' (or some symbol for 'The'). The result is (1*), for which (1) is shorthand.

Thus: (1*) *does not contain a symbol (simple or complex) that has simply been substituted for 'The F'*.

In the case of what Russell called a *logically proper name*, sentences containing the name are fully meaningful only if the bearer of the name exists. So if 'a' is a logically proper name, then 'Fa' is not meaningful if 'a' lacks a referent. It is clear, on Russell's analysis, that it is not a presupposition of the meaningfulness of (1) that the F exists. If the F does not exist, then (1) is false and therefore not meaningless. This is obvious, since the meaning of (1) is given more explicitly by (1*). *This is the point of saying that definite descriptions are not referring expressions* (or that they 'have no meaning in isolation'), and this is why Russell says that the relation between 'The F' and the F, when the F exists, should not be called the relation of reference or meaning; he calls it the relation of 'denoting' (but we will not follow him in that point of usage).

Note finally that since 'The King of France is wise' is *false* on Russell's theory, its *negation* is true. That is, 'It is not the case that the King of France is wise' is true. Thus for Russell, we have to distinguish the **internal** from the **external negation** of 'The F is G'.

The F is not-G (internal).

means ‘There is exactly one F and it is not G’, whereas

Not: the F is G (external).

means merely that it is false that the F is G. An *internal* negation can be false when the corresponding external negation is true. Consider the following example:

The Loch Ness Monster is not swimming. (internal, false)

Not: the Loch Ness Monster is swimming. (external, true)

The first requires for its truth the existence of the monster who is not swimming, whereas the second requires *either* that the monster exists but is not swimming *or* that the monster simply doesn’t exist. Internal negations always entail their corresponding external negations but not vice versa.

The distinction has to do with the position of the negation sign. In a bit of terminology that is often used: in the internal negation, the negation sign is *within* the **scope** of the definite description (which takes **wide** scope with respect to negation). In the external negation, the negation sign is *outside* the scope of the description (which takes **narrow** scope with respect to negation).

Spelled out quasi-symbolically, these are:

Internal negation:

There is an x such that (Fx & for every y (Fy \rightarrow y = x) & not: Gx)

External negation:

Not: There is an x such that (Fx & for every y (Fy \rightarrow y = x) & Gx)

● APPLYING THE THEORY OF DESCRIPTIONS

I–II deal with existence issues; III–IV deal with identity issues.

I In Frege’s theory,

(3) The man who killed the Kennedys is Cuban.

is not true and not false either: it is meaningful but neither-true-nor-false. The reason as we saw in Chapter 2 is that for Frege, this sentence is of the form

Fa.

It ascribes the predicate ‘is Cuban’ to the referent of the singular term ‘the man who killed the Kennedys’; the sentence is true if the predicate applies to that thing and false if it does not apply to that thing. That is, it is false if and only if its negation is

true. Since the singular term lacks a referent, this procedure never gets started, so no truth-value is determined. In Russell's theory, as explained above, the sentence is false; in Russell's theory, (3) is *not* of the form Fa , and, in a strict sense, 'the man who killed the Kennedys' is not a singular term at all.

II Consider this sentence (and assume that 'the Loch Ness monster' is a definite description, meaning 'the monster who lives in Loch Ness'):

(4) The Loch Ness monster does not exist.

As explained in Chapter 2, there are two ways to interpret the word 'exists'. First, we can take it as an ordinary first-level predicate, like 'is tall' or 'is wise'. In that case, it is true of every object, and not false of any object. Thus the 'nonexistence' predicate 'does not exist' is false of every object. In that case, how could (4) be true? Take again Frege's theory. To determine the truth-value of (4), we must first find the referent of 'the Loch Ness monster' and ask whether the nonexistence predicate is true of it. Suppose 'the Loch Ness monster' does have a referent. Then it exists, so 'the Loch Ness monster does not exist' is false. So if (4) were false, then Frege's theory would have no problem. But (4) is true. How can Frege's theory determine it to be true? If 'the Loch Ness monster' has no referent, then (4) comes out *neither-true-nor-false*, and *not* true – for the same reason that (3) came out that way on Frege's theory.

On Russell's theory, (4) comes out true, if we take 'exists' to be an ordinary first-level predicate. However, neither Frege nor Russell take that view of 'exists'. They both take existence to be expressed by the existential quantifier. We can attach it to a predicate like 'is a dog' to say that dogs exist – 'There is an x such that (x is a dog)' – but it cannot meaningfully be attached to a singular term: that is, we cannot meaningfully write 'For some x (Fido)'; it is not a sentence. What, then, of (4)? The trouble for Frege is that he has no direct way to express the content of (4): if 'the Loch Ness monster' is a singular term, then (4) simply doesn't contain a predicate to which the existential quantifier might be attached to become a sentence which can then be negated.

The difficulty is not ameliorated if we suppose that the real form of (4) is

(4*) Not: There is x such that (x = The Loch Ness monster).

This does now contain a predicate: ' β = The Loch Ness monster'. This predicate appears in the context of an identity statement which is generalised on its left side and negated. Identity statements are true if the object named on the left side is the self-same object as the object named on the right side, and false if the object named on the left side is a different object from the object named on the right side. The statement as a whole says that for every object taken as the left-hand object x , the identity statement ' x = the Loch Ness monster' is false. But there is no object to play the role of the right-hand object. The problem lies with Frege's theory of the semantical functioning of terms such as 'The Loch Ness monster' as referring terms.

For Russell, (4) presents no problem, since definite descriptions already contain quantifiers. If we take *F* as meaning 'is a monster who lives in Loch Ness', then (4) comes out simply as:

(4**) Not: There is an *x* such that (*Fx* & for all *y* (*Fy* → *y=x*)).

(This would also be true if there were more than one Loch Ness monster; but this is arguably the right result, for in that case it would not strictly be correct to speak of 'The' Loch Ness monster).

We said in the last chapter that the problem of negative singular existentials does not arise for negative existentials in general – that it arises only for negative *singular* existentials. Frege and Russell agree that, for example, 'Dragons do not exist' is equivalent to 'For every *x*, *x* is not a dragon', which is easily symbolised as 'For every *x*, not: *Fx*', or 'Not: There is an *x* such that *Fx*'. These sentences are obviously true (as is verified by checking each object for dragonhood).

III How does the theory of descriptions solve the problems of identity? From Russell's example of George IV, we can write:

(5) George IV wondered whether Scott = the author of *Waverley*.

(6) It is not the case that George IV wondered whether Scott = Scott.

Both of these are true. If we assume that 'the author of *Waverley*' is a referring expression, then from (5) and

(7) Scott = the author of *Waverley*.

which is true, we could infer

(8) George IV wondered whether Scott = Scott.

which contradicts (6). That is, (7) seems to tell us that we can substitute 'the author of *Waverley*' for 'Scott' wherever 'Scott' appears, in which case we can move from (5) to (8). The difficulty is evaded by the theory of descriptions in the following way. The theory of descriptions tells us to recast

(9) Scott = the author of *Waverley*.

as

(10) There is an *x* such that (*x* wrote *Waverley* & for every *y* (*y* wrote *Waverley* → *y=x*) & *x=Scott*).

(This is still of the form 'The *F* is *G*', where '*Fβ*' is '*β* wrote *Waverley*' and '*β* is *G*' is '*β* = Scott'). Thus (5) becomes

(5*) George IV wondered whether: there is an x such that (x wrote *Waverley* and for every y (y wrote *Waverley* $\rightarrow y=x$) and $x=Scott$).

(5*) does not appear to contradict (6). Notice that the only singular term (besides 'George IV' and '*Waverley*') in (5*) is 'Scott', so we cannot use (9) to get (8) from (5). The definite description 'disappears on analysis'; what is revealed is that, despite initial appearances, there is no singular term for which 'Scott' can be substituted.

There is a final complication. Suppose:

(11) Mary does not believe that every ghost is a goblin.

We might reason: since neither ghosts nor goblins exist, the two predicates are co-extensive, and each can be substituted for the other. So it looks as if we'd have 'Mary does not believe that every ghost is a ghost'! In other words, if Russell accepts that predicates with the same extension are co-referential (as Frege did), then it seems that he's got trouble. But he doesn't: he allows that co-extensive predicates may refer to different entities (in particular, to 'universals' – i.e. properties and relations). We'll see more of this doctrine below.

• NAMES AS DISGUISED DEFINITE DESCRIPTIONS

Russell was well aware that existence and identity problems arise for ordinary proper names as well as for definite descriptions. For example,

(12) Pegasus does not exist.

is seemingly both meaningful and true, and the following triad is seemingly consistent:

(13) John believes that George Eliot is an author.

(14) It is not the case that John believes that Marian Evans is an author.

(15) George Eliot = Marian Evans.

Russell's response is to suppose that ordinary proper names are not *really* referring expressions but *disguised definite descriptions*. For example, the meaning of 'Pegasus' is really the same as that of 'the winged horse', or some such thing. So far, this move reminds one of Frege, but with Russell's theory of descriptions, the meaning *and truth-value* of (12) can be explained, whereas with Frege's theory, as we learned in Chapter 2, they could not be. Frege's theory has it that (12) is meaningful but that it is neither true nor false – not, as according to Russell's theory, that it is true, which is the correct answer.

What about 'George Eliot', then? In fact, it is implausible to say that there is precisely *one* definite description that gives the meaning of that name. For example, A might know George Eliot only as the author of *Middlemarch* and *Daniel Deronda*, when B knows her only as the author of *Silas Marner* and *The Mill on the Floss*. Yet

surely both A and B may be credited with understanding such a sentence as ‘George Eliot was interested in German philosophy’. Accordingly Russell does not claim that each ordinary proper name is equivalent to some *one* definite description. Rather, he claims that on each occasion whereby one *uses* or *understands* an ordinary proper name, the name is equivalent in the person’s thought at that moment to a definite description. Thus for any sentence ‘... *a* ...’ where *a* is a proper name, the claim is:

For any subject B, any proper name *a*, and any occasion whereby B understands or meaningfully uses *a* in a sentence ‘... *a* ...’, there is a description ‘The F’ such that ‘... *a* ...’ is synonymous for B on that occasion to ‘... the F ...’

Thus the idiosyncrasy of proper names: each *use* of an ordinary proper name serves to abbreviate a definite description, but typically a name will not abbreviate the same description whenever it is used.

● KNOWLEDGE BY ACQUAINTANCE AND KNOWLEDGE BY DESCRIPTION

Suppose I know a man called Mr Wiggins only as the village baker and do not know that he has a daughter. Suppose you know the self-same Mr Wiggins only as the father of your friend Esmeralda but you do not know about the business he is in. Neither of us has met him face to face or has any other identifying information about him. According to Frege, if I say to you ‘Mr Wiggins is ridiculous’, you, knowing Mr Wiggins, will respond as if you understand my statement. But I will understand the statement as ‘The village baker is ridiculous’, and you will understand the statement as ‘Esmeralda’s father is ridiculous’. But, then, what is the basis for saying that we communicate? There is no proposition that we both understand.

Frege responds to this apparent difficulty by saying that for ordinary, non-scientific purposes, we simply tolerate such cases of non-overlapping knowledge of the referents of proper names. So long as the referent is the same, then normally we don’t get into trouble, and we can say that we communicate in a kind of downgraded sense of ‘communicate’.

Russell has more or less the same difficulty but has a more painstaking response. Suppose we ask: if ordinary proper names actually abbreviate definite descriptions, then are *any* singular terms genuine referring expressions? Are there any *logically proper names*, as Russell puts it?

It would be barely tolerable if there were not. For in that case, there would be no atomic sentences at all! And there had better be. For suppose there were not. Then all sentences would be generalisations – that is, sentences of the form:

For every *x* (... *x* ...)

There is an *x* such that (... *x* ...).

(or truth-functional compounds of these, but set those aside). What does it mean to say that every x is F ? It means that for each object x , Fx is true. That is to say, our understanding of a generalisation 'For every x , ... x ...' is by virtue of our understanding that it is true just in case *each instance* of it is true. 'For every x , Fx ' means ' Fa and Fb and Fc and ...', where a , b , c and so on, are all the objects that exist. So our understanding of generalisations seems to be based upon our understanding of their instances.

What was the reason for supposing that ordinary proper names are not logically proper names? There were two reasons. First, the existence of a bearer of the name is not a presupposition of the meaningfulness of a sentence containing the name. Second, if n is an ordinary proper name and n^* some other term such that $n = n^*$ is true, n is not thereby interchangeable with n^* in all sentences *salva veritate* (that is, without changing the truth-value of the sentence). Substitution failure occurs, for example, in sentences about propositional attitudes: John might believe that Hesperus is a planet but not that Phosphorus is a planet, despite the truth of 'Hesperus = Phosphorus' and so on.

How do we *know* that a sentence containing an empty proper name, like 'Pegasus flies', is meaningful? It seems that we know this because we recognise that someone could take an attitude towards the content of 'Pegasus flies'. It could be true that John believes that Pegasus flies. Since belief is an attitude towards a proposition, there must be such a proposition as that Pegasus flies. So the reason that 'Pegasus' is meaningful yet empty is that it is possible that one could be mistaken as to whether Pegasus has a bearer. So plausibly: a *logically proper name* must be a name such that one *could not be mistaken* as to whether it has a referent.

Similarly, how do we know that, despite the truth of 'George Eliot = Marian Evans', 'George Eliot' and 'Marian Evans' do not have the same meaning? Answer: because someone could reasonably disbelieve that George Eliot = Marian Evans. So plausibly, a logically proper name must be a name such that one *could not be mistaken* as to which object is its bearer.

Putting these points together: logically proper names are such that one could not be mistaken as to the *existence* of its bearer, and one could not be mistaken as to the *identity* of its bearer. The reasons that Russell cites for denying that such-and-such term is a logically proper name will no longer be in force.

Are there such names? Yes. For there are things about whose existence and identity it seems we are never mistaken (or so Russell thought). These things include:

- I Sense-data
- II Universals
- III The self

We will ignore III, as Russell was himself unsure, and it raises complex issues that go too far beyond the philosophy of language. As for the other two:

SENSE DATA

If you are experiencing a patch of red in your visual field, or a tickle in your throat, then it seems literally absurd to suppose you might be mistaken, or mistakenly suppose the sense-datum to be identical to another. You cannot mistake a sound for the colour red. For such things, ‘to exist is to be perceived’, as Berkeley put it; ‘to be this one rather than that one is to be perceived to be this one rather than that one’, he might have added. There is also a ‘Cartesian’ angle: plausibly, one is infallible concerning the contents of one’s mind, of one’s consciousness.

Now, of course, we don’t give names to our sense data like ‘Sigmund’ and ‘Elizabeth’. Russell’s idea is really this. We do, of course, in a certain sense talk about things in the world, such as tables. Suppose then that I see my cluttered table and say ‘The table is cluttered’. According to Russell, what I really do in such a case is to talk about the table *via* my sense data of it. I could conceivably be mistaken as to the existence of the table but not about the sense impression I am having. So what I do is to think:

The external cause of THIS (table-shaped batch of sense-data) is cluttered.

I don’t say the word ‘this’, but, so to speak, I use it in thought to refer *directly* to my sense data; so I am using it as a logically proper name. But the table itself is only ‘denoted’ in Russell’s sense (by the definite description ‘the external cause of THIS’) and not referred to directly. And, of course, the definite description is to be analysed in the Russellian way.

UNIVERSALS

Suppose I think that the Earth is round. (Pretend for simplicity that ‘the Earth’ is a logically proper name.) Of course, I can imagine being mistaken as to whether there are any objects that instantiate the property; there are surely properties about which human beings have been mistaken in that way. But that is not the issue; could I be mistaken as to the very existence of the property roundness? It hardly seems possible. Could I be mistaken as to whether or not roundness is identical with some other property? Well, this might be imaginable. For example, perhaps someone might think that *being equiangular* and *being equilateral* are the same property when they are not. And one might be mistaken as to which property is *called* roundness. But set these cases aside: let us assume that when we refer to or mean a property such as roundness in such a proposition as *The Earth is round*, we think of the property of roundness in such a way that we could not be mistaken as to either the existence or identity of the property. Likewise for any universal, as it enters into thought when we think a thought whose content is an atomic proposition.

As Russell puts it, the proper *subjects* of an atomic proposition are sense-data, the entities that we *perceive*. The entities playing the role of the predicates of an atomic

proposition are universals, the entities that we *conceive*. Universals are creatures of thought, sense-data creatures of perception. This is reflected logically in their differing roles in the structure of atomic propositions. These are things with which we have a special epistemic relation – to wit, that we cannot be mistaken about either their existence or identity. Russell calls this the relation of **acquaintance**, and it leads him to formulate his famous

Principle of acquaintance: any proposition we can understand must be composed entirely of entities with which we are acquainted.

We use such knowledge of items of acquaintance to form propositions. Since we are not strictly speaking acquainted with tables and trees, it follows that we cannot understand atomic propositions about, or which ‘contain’, tables and trees. But how then do we talk about, and communicate about, such things?

Russell’s answer is somewhat perplexing but not obviously incorrect. Suppose there is a table before us, and I say ‘The table is round’. I denote the table *via* a description containing reference to my sense data and the notion ‘the external cause of ...’ Call this description ‘the external cause of ϕ ’, where ϕ is the relevant batch of sense-data. You understand what I say by means of your own sense-perceptions, using some description ‘the external cause of ψ ’. But if all goes well, the external cause of ϕ = the external cause of ψ . So the situation is:

I think: the external cause of ϕ is round.

You think: the external cause of ψ is round.

The external cause of ϕ = the external cause of ψ .

Neither of us is equipped to think, to apprehend the thought, that the external cause of ϕ = the external cause of ψ , since we’re not acquainted with each other’s sense-data. And we lack what Russell calls **knowledge by acquaintance** of the table. But – scepticism aside – since I know that the external cause of ϕ exists, and you know that the external cause of ψ exists, we both know that *there is* an object x (the table) by virtue of which the thing you are thinking of is the thing I am thinking of. We both have **knowledge by description** that there is such a table. Thus we both know that there is a **singular proposition**

a is round

which is equivalent in this context to the proposition that I think and to the one that you think. Yet neither of us can grasp this proposition (because neither of us is directly acquainted with a , but we both know that there is such a proposition as *a is round*).

The resulting picture of the relation of empirical thought to reality is very similar to that of Locke and perhaps Hume. Largely for that reason, Russell is generally regarded as an empiricist.

● HISTORICAL NOTES

We have considered only a narrow time-slice of Russell, though arguably it is the time of his greatest accomplishments. In his earlier book *The Principles of Mathematics* (1903), Russell thought of language as more-or-less transparent, that one could ‘see’ through it to the entities meant, and that the logical form underneath – of facts, universals and objects, etc. – corresponded closely with the outer form of its linguistic expression. He subscribed to many elements of naïve semantics – with the clearest departure from it represented by his theory of ‘denoting concepts’, precursor to the view espoused in ‘On Denoting’ (1905) and which in some ways aligned him with Frege. The view encouraged him to think that questions of language could often be ignored; one is talking directly about the subject-matter of language, not language itself. That view gradually changed, most decisively with the 1905 theory of descriptions, according to which logical form may be very different from grammatical form. Since Frege was largely unknown, analytical philosophy is often said to date from this article, because Russell discovered the vein, as it were, which logically trained philosophers are uniquely equipped to mine: beneath the hubbub of ordinary language, there lies hidden the crystalline realm of propositions in their purity. The time-slice of Russell that we have presented here is the Russell of his book *The Problems of Philosophy* (1912; it was also the time when he wrote with A. N. Whitehead the magnificent *Principia Mathematica* – 1910, 1912 and 1913 – which developed the thesis that mathematics is founded on logic). That view rapidly gave way to his ‘logical atomism’ (1918, 1924), according to which, very roughly, ordinary material objects are identified with classes of sense-data and are not the causes of them. The young Ludwig Wittgenstein, who began studying with Russell in 1911, came out with his *Tractatus Logico-Philosophicus* in 1921; and Rudolf Carnap delivered a rigorous development of a view that is in various ways similar to Russell and Wittgenstein’s logical atomism, in his *The Logical Structure of the World* (1928). Another figure who accepted the basics of Russell’s logical atomist view was Alfred Ayer; we’ll consider Ayer, Carnap and the young Wittgenstein in the next chapter. W. V. Quine, of whom we’ll hear much more in Chapters 6 and 11, was profoundly impressed by Russell’s work in the foundations of mathematics as well as by the theory of descriptions, but he also famously presented an alternative to the picture of knowledge, named ‘holism’, which differs markedly from Russell’s scheme of 1912 (but not so much from later iterations: Russell 1921, 1927, 1940, 1948). All of these figures, along with Frege, thought that existence is not a property of objects (like *roundness*), but rather is expressed by the existential quantifier. Alexius Meinong (1853–1920) famously held the contrary view and was often the opponent of Russell’s writing on the subject. ‘Meinongism’ has since become popular among today’s metaphysically inclined philosophers of language; it is considered briefly in Chapter 13.

● CHAPTER SUMMARY

Russell sets out to solve the puzzles addressed by Frege’s theory of sense and reference, without positing the category of sense in addition to reference. There are

fundamentally two such problems: that of nonexistence, and substitution within propositional attitude contexts.

First, Russell sets out his theory of definite descriptions. All complex singular terms can be written as 'The F', and – except for certain cases – any sentence containing 'The F' is of the form 'The F is G' for some (possibly complex) predicate '*a* is G'. 'The F is G' is equivalent to 'There is an *x* such that *Fx*, and there is not a *y* such that *Fy* and *x* ≠ *y*, and *Gx*' – which is 'There is one and only one thing such that it is F, and all F are G'.

Second, he claims that all ordinary proper names are really definite descriptions in disguise; they abbreviate them. For example, on a particular occasion of use, one's use of 'Bob' might mean, in sense of being strictly equivalent to, 'the man standing before me'.

Thus if '*a* is G' is analysed as of the form 'The F is G', then it is simply false if there is no such thing as *a* (Frege's account has it as neither-true-nor-false). Furthermore, if there is no such thing as *a*, '*a* does not exist' comes out as 'Not: There is exactly one *x* such that *Fx*', which is true; Frege's scheme runs into trouble in getting the right truth-value in such cases, called *negative singular existentials*.

The key to solving the puzzle of propositional attitudes is again to analyse definite descriptions in the same way, whether they are explicit or disguised; once analysed, Russell believes, the puzzle disappears. For example, granted that Scott = the author of *Waverley*, the conjunction 'George IV wondered whether Scott = the author of *Waverley*' and 'It is not the case that George IV wondered whether Scott = Scott' looks to be inconsistent. But the apparent inconsistency disappears when the definite description 'the author of *Waverley*' is analysed.

For Russell, the only genuine logically proper names – ones that cannot be analysed as disguised definite descriptions – are ones standing for sense-data (and possibly for the self). More generally, Russell propounds his famous distinction between knowledge by acquaintance and knowledge by description; the crucial principle is his principle of acquaintance: any proposition we can understand must be composed entirely of entities with which we are acquainted.

● STUDY QUESTIONS

- 1 Could Frege borrow Russell's strategy of distinguishing internal from external negations in order to cope with true negative singular existentials?
- 2 Use Russell's theory to analyse 'It is not the case that Nessie breathes air'. How does the distinction of scope apply?
- 3 Alexius Meinong held that existence *is* (more or less) a predicate – that such things as Pegasus or Santa Claus may fail to exist, but nevertheless they are *objects of thought*. Russell held something similar at one time but dropped it when he came upon the theory of descriptions. Was he right to? Isn't Meinong's theory better just because it does not involve a distinction between surface grammatical form and deeper logical form?

- 4 Is Russell's account satisfactory of how we communicate about ordinary things such as tables and trees?

● PRIMARY READING

Russell, B. (1905) 'On Denoting,' reprinted in his *Essays in Analysis* pp. 103–119, and in *Logic and Knowledge*, pp. 41–56.

Russell, B. (1911) 'Knowledge by Acquaintance and Knowledge by Description', reprinted in his *Mysticism and Logic and Other Essays*, pp. 152–167.

● SECONDARY READING

Hylton, P. (2005) *Propositions, Functions, and Analysis: Selected Essays on Russell's Philosophy*. Contains several essays that are accessible yet penetrating, including a comparison of Russell with Frege.

Neale, S. (1993) *Descriptions*. A lucid exploration of Russell's theory of descriptions.

Sainsbury, M. (1979) *Russell*. A broad introduction to the main aspects of Russell's philosophy.

● NOTES

- 1 At one point, many astronomers erroneously believed there to be a planet orbiting the sun within the orbit of Mercury and called it Vulcan; see Chapter 2, 'Two problems for naïve semantics'.
- 2 Actually examples such as 'Sally's teacher' do not imply uniqueness in the way that 'Sally's father' does. Indeed, it is arguable that the uniqueness of Sally's father is not determined by semantics but by the facts of biology, and 'Sally's teacher' in the context 'Sally's teacher was talking' is equivalent to 'A teacher of Sally's', hence not a singular term. We can think of such expressions as 'Sally's teacher' or 'Sally's father' to be ambiguous, implicitly awaiting transformation with a quantifier into forms such as 'The teacher of Sally', 'A teacher of Sally' or 'Any teacher of Sally'. Normally the context clarifies which quantifier is meant.
- 3 This also is more complicated; we ignore apparent counterexamples involving modal vocabulary and the propositional attitudes.
- 4 In fact, it is equivalent to the yet more compact: "There is an x such that for all y , $((Fy \leftrightarrow x=y) \ \& \ Gx)$ ", or ' $\exists x \forall y ((Fy \leftrightarrow x=y) \ \& \ Gx)$ '. A nice exercise in logic is to derive the equivalence of the two.

4

• russell's theory of judgement, the early Wittgenstein and logical positivism

Frege and Russell both came to their ideas in the philosophy of language via their work in logic. Frege more than Russell is credited with the invention or discovery of the logic of truth-functions and quantifiers that is studied in elementary logic courses to this day. The new logic went far beyond the extant logic of the syllogism initiated by Aristotle; it dealt not only with simple inferences involving categories – one-place predicates – but also with the logic of whole propositions and relations of arbitrary complexity.

Their interest in logic, in turn, came via their interest in the foundations of mathematics. They sought to establish beyond doubt the thesis of **logicism**: that pure mathematics reduces to pure logic. Frege had completed the crucial Volume I of his *Grundgesetze der Arithmetik* in 1893, but his work was largely unknown to Russell until his own *Principles of Mathematics* (1903) was near completion. In the Appendix, added just before the book went to press, Russell expresses no little admiration for Frege, but does point out – as he did in a famous letter to Frege of the same year – that in Frege's system a contradiction is derivable, a difficulty which has since come to be known as Russell's paradox. Frege's system – in particular his Axiom V – implies that for any predicate, there is a set or extension corresponding to it (in other terms, that every condition determines a class).¹ It follows that some sets are members of themselves: for example, the set of sets that have more than three members. Other sets are not members of themselves: the set of leopards, for example, is not itself a member of the set of leopards (that is, the set of leopards is not itself a

leopard). Consider then the predicate 'is not a member of itself'. This too has a set corresponding to it, according to the implication of Axiom V: the set of sets which are not members of themselves. Call this the 'Russell set'. Is the Russell set a member of itself, or not? If the Russell set is a member of itself, then it is *not* a member of itself. If the Russell set is not a member of itself, then it *is* a member of itself. That is the contradiction: a proposition is either true or false, but not both; in this case, if it is true, it is false, and if it is false, it is true. Something is wrong in Frege's system.

Frege replied to Russell that the contradiction 'seems to undermine ... the only possible foundations of arithmetic as such'. Indeed, he never was successful in repairing the defect. But Russell, together with Alfred North Whitehead, came out with the colossal work *Principia Mathematica* in three volumes in 1910, 1912 and 1913 (the second edition appeared in 1927). The work was much more complicated than Frege's, but plausibly made the case for logicism, that logic is the real subject-matter of pure mathematics.

Frege and Russell both saw that with their new, vastly more powerful logic, it was possible to extend these thought-sharpening techniques to other sciences and to empirical domains generally. It was Russell who took concrete steps in this direction. In such works as *The Problems of Philosophy*, *Our Knowledge of the External World*, *The Philosophy of Logical Atomism* and the extended essay 'Logical Atomism', he attempted to establish that all knowledge is founded on knowledge of logic plus sense-experience. And this required analysis of the sentences in which knowledge-claims are housed. What objects are such claims really about? What are the properties or relations to which such claims are really committed?

In 1911 – perhaps at the suggestion of none other than Frege, whom he had visited earlier that year – the twenty-two-year-old Viennese Ludwig Wittgenstein went to Cambridge seeking Russell. It was not long before Russell thought him 'the most perfect example I have ever known of a genius as traditionally conceived; passionate, profound, intense and dominating'. And it is an aspect of Wittgenstein's early ideas – as expressed in his *Tractatus Logico-Philosophicus* of 1921 – that forms the fulcrum of this chapter. It is probably inaccurate to speak of the *Tractatus* in same breath as the famous movement known as Logical Positivism, but it is not inaccurate to characterise them both as receiving a significant stimulus from Russell's views immediately after 1910. One central idea of the *Tractatus* – which has since been called the 'picture theory' of meaning – can fruitfully be seen as a response to certain shortcomings of Russell's view. We will not stop to consider a sophisticated version of Russell's view, one that might perhaps overcome these shortcomings: we'll see just enough to understand why the picture theory might be thought to be superior.

● PROPOSITIONS, FACTS AND RUSSELL'S THEORY OF JUDGEMENT

Truth and falsity, we have been assuming, are features of propositions. It is natural to suppose that truth is *correspondence*, in particular that it is correspondence of a proposition with *the way the world is, the way that things are*. A way that things are is

a *fact* or *state of affairs*. There is, for example, the fact that snow is white, that Arsenal won the English Premiership title in 2002, that $2 + 2 = 4$. We intuitively suppose that facts are configurations of entities – for example, the fact that snow is white is composed of snow and the property whiteness. Singular terms that stand for this fact include: 'snow's having whiteness', 'snow's being white' or 'the fact that snow is white'.

Thus we might have tried in Chapter 1 the following characterisation of truth for atomic propositions:

Naïve definition of truth for atomic propositions*. An atomic proposition is *true* if and only if: (i) it is a one-place atomic proposition containing an object *o* and property *P*, and the fact that *o* has *P* exists; or (ii) it is a two-place atomic proposition containing objects *o*₁ and *o*₂, relation *R* and the fact that $R(o_1, o_2)$ exists (and so on for three-place relations, four-place relations etc.).

But Russell notes that there is a serious problem with this idea (hence the asterisk; see Chapter 1, 'Truth and meaning for atomic sentences'). Consider the *fact* that Mars orbits the sun. Compare this with the *proposition* that Mars orbits the sun. They are *both* composed entirely of Mars, the relation *α orbits β* and the sun. So what is the difference between them? They have to be different in some way, or they would be identical: the very self-same thing. In which case *every atomic proposition would be true* – no atomic proposition could be false since there would be a fact to which it corresponds, namely itself. Since it's thinkable, and therefore the proposition exists, that walrus fly, it would follow that it is true that walrus fly, which, presumably, they do not. The model seems not to allow for false beliefs, for false propositions.²

There is also a problem with the 'structural' conception of propositions outlined in Chapter 1, which at the time we merely glossed over. We said that the proposition expressed by 'Mars orbits the sun' contains, or is made up of, Mars, *α orbits β*, and the sun. But we did not say how this differs from the proposition expressed by 'The sun orbits Mars', which it surely does. Those two propositions assemble the three constituents in different ways: the *direction* of the relation is different in the two cases. Although it is clear that the difference in meaning is indicated by the differing linear order of the constituents of the *sentences*, we have not explained what, in *propositions*, reflects this difference.

Compare this situation with the notion of a *set* or *class*. Like propositions, sets are abstract objects that directly involve their constituents. But unlike the case of propositions, a set is defined by its constituents, its members – it is necessary and sufficient for A and B to be the very same set that they contain just the same members. So {Pluto, the relation *α orbits β*, Venus} is the same set as {Venus, the relation *α orbits β*, Pluto}. Whereas, as we just saw in the case of the propositions expressed by 'Mars orbits the sun' and 'The sun orbits Mars', two distinct propositions may have the same parts or constituents.

Closely related to this is a difficulty we encountered in connection with the concept horse problem in Frege, discussed in Chapter 2, 'Predicate reference and the concept

horse problem'. Not every collection of referring expressions coheres meaningfully in such a way as to express a proposition. 'Socrates' and 'Groucho Marx' do not together make up a meaningful sentence. At most they constitute a list, perhaps a list of great raconteurs. But further, consider 'Socrates' and 'the property of being a drinker', or 'Socrates', 'drinks'. There is a difference between the proposition that *Socrates drinks* and the list *Socrates, drinks*. What is the difference? It does not help matters if we suppose that 'drinks' is shorthand for 'is a drinker' or 'is drinking', thereby importing the copula, the 'is' of predication, into the scene. For what is imported has not to be a mere sound or squiggle on paper; it must be a meaningful entity. The thing that it means must be a kind of abstract entity which we can call *copulation*. But then we are back where we started, with no way of distinguishing the proposition *Socrates is drinking* from the list *Socrates, copulation, drinking*.

In view of these problems, Russell introduces the **Multiple Relation Theory** of judgement. Forget propositions, Russell now thinks. Think of truth not as a feature of propositions, but of those mental states or facts called judgements or beliefs. Yes, if there were no judgers, then there would not be anything true (or false). But there would still be *facts*. The only reason to dislike facts was the difficulty of admitting them in *addition* to propositions, but the strategy now is to get rid of the latter.

Crucially, Russell relies on a notion we've seen before: the relation of *acquaintance* (Chapter 3, 'Knowledge by acquaintance and knowledge by description'). We stressed that, aside from one's self, the only *objects* with which one can be acquainted are sense-data, but that will not matter for present purposes. For simplicity, we can assume that one can be acquainted with any ordinary object. Also, one can be acquainted with many entities which are not objects (in another terminology, which are not *individuals*); in particular one is acquainted with **universals** (properties, relations). As Russell puts it, objects are *perceived*; universals are *conceived*.

Suppose then that a judger B is going to judge that *Rab* ('The cat is on the mat', etc.). **First**, B must be *acquainted* with (must refer to, must think 'of') the objects a, b and the relation R (Figure 4.1).

The dotted lines indicate acquaintance.

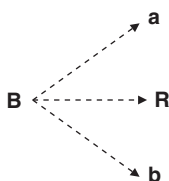


Figure 4.1 Acquaintance

Second, B forms a certain attitude towards this complex: B judges that *Rab* (indicated by the solid lines) (Figure 4.2).

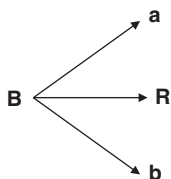


Figure 4.2 Judgement

The judgement that *Rab* is a *four-termed* relation between the judger B and a, b and R. The judgement is itself a fact. The judgement is true if there is a certain fact *not* involving B, namely the fact that *Rab*. The judgement is false if there is no such fact. Truth is an objective correspondence between facts – in particular between *mental states* and reality, not between abstract *propositions* and reality.

There is more complexity to the theory, but that will do for getting across the main idea. Russell hoped that propositions had been explained away as a particular type of mental fact, and that truth and falsity had been explained as attributes of those mental facts.

• THE *TRACTATUS LOGICO-PHILOSOPHICUS*

There are many, many aspects of the *Tractatus* that we shall not go into, fascinating though they are. We want to get a basic handle on its 'Picture Theory' of the proposition, and to seize a few of the main features that inspired later philosophers: in particular the Logical Positivists or Logical Empiricists.

In 1912, when Wittgenstein was 22 years old, he wrote in a letter to Russell:

If I judge that *aRb*, I want a theory of judgement to tell me *aRb* or $\sim aRb$ *without any further premise*. The condition is not fulfilled by your theory.

(For 'Rab' Wittgenstein follows Russell in writing '*aRb*', and for 'not *aRb*' he writes ' $\sim aRb$ '.) What he wants a theory of judgement to do can be displayed by the following inference:

(1) B judges *aRb*.

Therefore

(2) *aRb* or $\sim aRb$.

Why can we have (1) but not (2)? The problem is that Russell's theory of judgement assumes the relation of *acquaintance* as basic, and thereby seems to assume

that there is no constraint of the right kind on the entities with which one may be acquainted. I can perceive, be acquainted with, the table. Likewise with the penholder. Likewise with the book. I can be acquainted with all three – indeed, with all three at one time. What then, Wittgenstein asks, is to stop me from taking a judging attitude towards them; from judging that *the table penholder the book*?

To cut a long story short, nothing in Russell's theory prevents it, thinks Wittgenstein. So far as the Multiple Relation Theory is concerned, the objects of judgement could fail to cohere in the right way. As we can hardly help saying, it could fail to be a proposition. But propositions or thoughts are, if nothing else, the contents of judgement. That much is non-negotiable. It is not adequate to reply that it's just that *the table penholder the book* could never be true, that is, could never be a fact. The problem is not that for all that's been said the judgement might be true when it could not be true, but that there could be such a judgement at all.

The picture theory. I will run roughshod over Wittgenstein's terminology, for the sake of getting a rough presentation of the picture theory before us.

Wittgenstein agrees with Russell on one basic point: truth must be explained in terms of facts. But he does not agree that to speak of propositions is a mere *façon de parler*. He thinks that there are such things as propositions. They are not objects, but are themselves facts (the enumeration is that of the *Tractatus*):

3.14 What constitutes a propositional sign is that in it, its elements (the words) stand in a determinate relation to one another. A propositional sign is a fact.

Propositions are not abstract entities *expressed* by sentences; rather, they simply *are* meaningful declarative sentences (in the context of utterance). They are *types* of ink-marks and vocalized phonemes, etc., namely ones that are meaningful in the context.

Thus take an everyday fact, an everyday state of affairs, such as the cat's being on the mat. The cat-on-mat is not an object; it is a fact, and facts are not objects (we will see why shortly). Consider now a simple, accurate, ordinary picture or drawing of that fact. The picturing relation is a relation between two facts: the picturing fact and the pictured fact. The picturing fact must have an item which stands for the cat and an item which stands for the mat, and must show how the one is related to the other.

Now consider the *proposition* (the sentence) about the same cat and the same mat as before:

The cat is on the mat.

Like the picture, an item – 'the cat' – stands for or refers to the cat; an item – 'the mat' – stands for or refers to the mat, and the proposition *shows how* they are related. Except in this case the 'showing' is more abstract, and depends explicitly on convention. In the actual, physically realised picture, it is accomplished by embodying the relation of one thing's being on another, in that one set of marks is spatially immediately above another set of marks. (That that is so emerges even more clearly from a case, which Wittgenstein cited as an inspiration, of toy cars used in a courtroom to

depict a traffic accident.) In the proposition, we take advantage of the convention of English that writing one name and another on each side of the expression 'is on' shows that the one to the left is atop the one to the right. The propositional sign says that one thing is on top of another: it *shows how they must be related if the proposition is true*.

Thus we get formulations such as the wonderfully succinct 3.1432:

Instead of, 'The complex sign "aRb" says that a stands to b in the relation R', we ought to put 'That "a" stands to "b" in a certain relation says that aRb.'

English could well have been different. It might have been such that

The cat
the mat.

was our way of writing what we actually write as 'The cat is on the mat'. Or the conventions of English might have directed us to represent the very same state of affairs by writing 'The mat is on the cat' (it is perhaps useful to note that Arabic and Hebrew are written from right to left, or that traditional Japanese is written top to bottom).

Thus Wittgenstein says:

3.1431 The essence of a propositional sign is very clearly seen if we imagine one composed of spatial objects (such as tables, chairs, and books) instead of written signs.

It is clear that if, as it happens, the cat is *not* on the mat, the status of the spatial fact as a picture is not affected; it is simply an inaccurate picture, a false picture; things are not as the picture indicates – similarly with the proposition, the purely 'logical' picture. The view has no problem allowing for false propositions.

The picture theory purports to solve the list problem. The problem with 'the table penholder the book', according to the present rules of English, is that it is not a sentence and indeed not a fact with the requisite structure to be a picture (think of the three items drawn individually on a page, but without their being unified into a single picture). The problem about distinguishing the 'direction' of a relation – *the cat is on the mat* from *the mat is on the cat* – is purportedly solved because the two sentences in English, as a matter of convention, say opposed things: Whereas the former shows the one atop the other, the latter shows the other atop the one.

The limits of sense. 'The world is all that is the case', Wittgenstein announces in the very first statement of the *Tractatus*. The world is all the facts, including ones that happen to be propositions. A fact is one or more objects having a relation or bearing a property. Suppose that $p_1, p_2, \dots p_n \dots$ are all the elementary propositions (true and false), corresponding to all possible elementary facts $f_1, f_2, \dots f_n \dots$ (whether or not they obtain).³ The whole of reality must then in principle consist in some

selection of all those elementary facts that happen to be the case, and so must be representable as the corresponding elementary propositions that happen to be true. Indeed, it is fundamental to the *Tractarian* world that no elementary proposition is *a priori* or necessary. Each is logically independent from the rest; its truth-value has no bearing on those of any other elementary proposition.

The picture theory is a theory of the senses or truth-conditions of *atomic* or *elementary* propositions; such a proposition is true if the pictured elements are as the picture shows them to be, and false if they are not. The sense or truth-conditions of a proposition *in general* – including non-atomic or complex propositions – is determined logically by those of the elementary propositions. There are two stages to appreciating this.

First, Wittgenstein gives an account of the sentential connectives (in virtually the same way as that in Chapter 1).⁴ Every proposition is true or false, not both. For any proposition p_l , the negation 'not- p_l ' is true if p_l is false, and false otherwise. For any two propositions p_l and p_k , their disjunction ' p_l or p_k ' is false if they are both false, and true otherwise. For any p_l , p_k , their conjunction ' p_l and p_k ' is true if they are both true, and false otherwise.

The same information can be represented diagrammatically in a truth-table (Table 4.1).

Table 4.1 Truth-table for 'not', 'or' and 'and'

p_l	p_k	not p_l	p_l or p_k	p_l and p_k
T	T	F	T	T
T	F	F	T	F
F	T	T	T	F
F	F	T	F	F

Examples combining these, such as ' p_l or (p_k and not- p_m)', can be explained along these lines (I am using parentheses in the expected manner but Wittgenstein's own notation differs; see note 4). The sense of a complex proposition composed of atomic propositions joined by these sentential connectives is its truth-condition – an exact representation of which is provided by its truth-table. These expressions – 'and', 'or' and 'not' – do not stand for any element in the realm of facts; they are not names. Again, reality itself is just some vast assemblage of atomic facts.

Second, Wittgenstein proposes a 'substitutional' account of quantification. An atomic proposition says that some object is a certain way (bears a certain property) or that some objects are a certain way (stand in some relation). Again, anything we wish to express by means of generalisation must already be present in those elementary propositions. If the objects are $o_1, o_2, \dots, o_n \dots$, named by ' $o_1, o_2, \dots, o_n \dots$ ', then 'every object is red' is equivalent to the (perhaps infinite) conjunction ' o_1 is red and

q_2 is red and ... and q_n is red ...' 'Some object is red' is equivalent to the disjunction ' q_1 is red or q_2 is red or ... or q_n is red ...' Quantifiers, much like the sentential connectives 'and', 'or' and 'not', are mere notational conveniences and do not stand for anything.

A note about the status of objects in the *Tractatus*. It has been useful to pretend otherwise, but for example a cat is not, for the *Tractarian* Wittgenstein, an object. For the cat can be disassembled, and for reasons rather similar to the reasons Russell provides (see Chapter 3, 'Knowledge by acquaintance and knowledge by description'), this shows that 'the cat' must be a description, not a (logically proper) name. The real objects must, for reasons whose exact nature will not detain us, be simple and indestructible; their existence guarantees the meaning of our words. They and their configurations are what language is ultimately about – it is where the 'feelers' of language 'touch' reality. It takes logical analysis to see how ordinary propositions reduce to propositions that are explicitly about these real objects. Wittgenstein is famously non-committal about precisely what the objects are – are they perhaps themselves universals? Or material atoms of some sort? He does not say (although he does fleetingly give a few examples, such as a colour patch or note of music, and he does recognise certain essential properties of objects: a note, for example, must have some particular pitch).

There are many problems, with both the theory and the foregoing representation of it, and I have rounded many corners, but the vision is important, however clouded it is. What Wittgenstein has provided is a model for all sense-making language. This means that any language that cannot be fitted to it in some manner fails to have sense.

It is possible to point to many sorts of example of that kind of failure and of propositions right at the limit:

- 1 Tautologies and contradictions. Sentences which are true in *all* conditions, such as 'p or not-p', cannot be said to say the world is any particular way. Similarly for sentences which are contradictory, which cannot possibly be true in *any* condition, such as 'p and not-p'. These are the limiting cases of sense.
- 2 Language which appears to explain the way that language works in general – for example, the picture theory itself. One language can be used to describe the conventions of another language, but logical form is what is common to all possible languages. It 'cannot itself be the subject of depiction. One cannot get away from it when depicting' (4.041). Logical form is ineffable: a step beyond any particular notation.
- 3 Language which expresses values, whether ethical, aesthetic, mystical or religious. This is language which is evidently not about contingent configurations of simple objects.

Now, Wittgenstein famously says 'in order to draw a limit to thinking, we should have to think both sides of the limit'. Nevertheless, *there is nothing wrong with this sort of nonsense* (or borderline cases as in variety #1). For language can *show* things which it does not literally *say*. When learning logic, or

the picture theory, or about values, one learns something but, strictly speaking, cannot say what one learns; one can only *show* what one learns. Language can show things of great importance, as well as confuse us when we mistake language for one when it ought to be taken for the other.

Wittgenstein's term for these varieties of nonsense is 'sinnlos'. There is another variety, called 'unsinnig' (somewhat similar to 'self-defeating'). Some but not all of these more serious varieties of nonsense are more dangerous in the way that they may mislead.

- 4 Less troublesome varieties of 'unsinnig' are:
 - a Contrary to the ordinary principles of grammar – 'Suzy that Jane it', 'The table penholder the book'.
 - b Apparently ok grammatically but lexicographically wanting, as in "Twas brillig, and the slithy toves did warble in the wabe'.
- 5 More troublesome varieties of 'unsinnig' are grammatically and lexicographically ok but in violation nonetheless according to the picture theory:
 - a 'Colourless green ideas sleep furiously' (Chomsky) or 'this stone is thinking about Vienna' (Carnap). What is wrong with these examples is something about the combination not making proper sense, not making for a *picture*, despite the grammatical well formedness. More subtle forms of this are more insidious.
 - b 'This sentence is not true'. The problem here is that there is no picture, no proposition, being said not to be true. 'This sentence' does not stand for an independent sense-making sentence.

For Wittgenstein, logic is not a science alongside astronomy and the rest. Logic is prior to all science, a precondition of thought. It is essentially ineffable because logical form is ineffable. So far as philosophy is concerned, he took himself 'to have found, on all essential points, the final solution of the problems'. But as his friend Frank Ramsey pointed out in his review of 1923, and as he came to accept some years later, he was mistaken in this. We turn now to a similar style of project which was however less absolutist, less monkish and more friendly to Russell: the Verificationist philosophies of A.J. Ayer and Rudolf Carnap.

● VERIFICATIONISM I: AYER

Ayer's *Language, Truth and Logic* appeared in 1936 and, like the *Tractatus*, was written when its author was still in his twenties, and, like Wittgenstein's book, suffers not from the doubts and hedges of an older philosopher. If it does not have quite the depth, originality and logical awareness of Wittgenstein's book, and is not possessed of the rigour and intellectual sophistication of Carnap's work of roughly the same period – the subject of the next section of this chapter and on which many of Ayer's main ideas are modelled – it remains the most vibrant and unflinching statement of classical verificationism, and is at least popularly the central text of the so-called 'Linguistic Turn' undergone by mid-century Anglo-American philosophy. Carnap is a more abstract,

elusive and difficult philosopher than Ayer, and their motivations differ markedly. Despite Carnap's having come first, I present the relatively down-to-earth Ayer first before sketching some Carnapian complexities. Although he came to disown much of it later in life, Ayer's book is very readable and scintillating at times. It looks back explicitly and satisfyingly to Berkeley and Hume, as well as looking forward to recent figures such as Michael Dummett, Crispin Wright and Simon Blackburn.

In the *Tractatus*, Wittgenstein characterised the limits of sense-making language in terms of the portrayal of facts, which are themselves contingent configurations of objects. Language that fails to 'picture' such configurations fails to say anything factual, although it may show something. There is no obvious *epistemological* constraint here; there is no palpable demand that in order to make sense one must speak only of what one knows or is acquainted with, or that one must in some less direct way be able to make mental contact with the subject-matter.

The showpiece of Ayer's book – his **Verification Principle** – makes explicit such a constraint, and together with its philosophical surroundings makes for a captivating picture of the nature and limits of language and philosophy. The philosophical surroundings include the equation of necessity and the *a priori* with the analytic, which, in turn, is equated with statements that are immune to confutation by experience because they 'simply record our determination to use symbols in a certain fashion' (Ayer offers various characterisations – tautologies, in Wittgenstein's sense, and 'formal truth' – without quite committing to any or making them exact; we'll follow him in speaking of 'analytic truth'). The denial of an analytic truth is self-contradictory: statements which are not analytic pertain to 'empirical matters of fact'; they are 'hypotheses – probable but never certain'. The exact reason for this will emerge presently.

The Verification Principle, to which all empirical hypotheses must correspond, is 'not ... that [a statement] should be conclusively verifiable,' but that 'some possible sense-experience should be relevant to the determination of its truth or falsehood'. This is cashed out further as *weak* verification: a statement is weakly verifiable 'if it is possible for experience to render it probable'. Ayer equates *strong* verificationism – that all empirical hypotheses must be *conclusively* verifiable – with the Positivism of the earlier Ernst Mach, which he rejects.

This makes it possible for Ayer to contend that all knowledge is either empirical or *a priori* but that the latter is merely 'linguistic in character' (57), by which he means not that such claims are themselves about matters of linguistics, but that the truth of such claims can be explained as analytic. He is thus more or less with Hume, who famously said we must 'commit ... to the flames' – statements which are not of either 'matters of fact' or 'relations of ideas'. Ayer is with Schlick and Carnap in espousing an anti-metaphysical philosophy, for metaphysics, in one sense, purports to be concerned with things beyond human experience, beyond the reach of empirical hypothesis, beyond verification by means of the senses. So the existence of a transcendent God, the possibility that sense-experience is unreal, the idea that necessarily unperceivable 'substances' underlie the qualities perceived in material objects, and many other things, are rejected – not as false or nonexistent, but because the sentences purporting to be about them lack sense. For example, Ayer says 'anyone who condemns the sensible world as world of mere appearance, as

opposed to reality, is saying something which, according to our criterion of significance, is literally nonsensical' (39).

Why, then, are *all* propositions which do meet the criterion of significance, unless they are analytic, merely probable? Ayer offers two steps to explain that surprising claim.

First, Ayer thinks of philosophy as logical clarification that it in no way competes with natural science. He follows Hume in thinking of the Principle of Induction – roughly, that if sufficiently many Fs have been observed and all have been found to be Gs, then it is probable that *all* Fs are Gs – as itself an empirical generalisation. It is not an analytic truth, and, indeed, we might be wrong in accepting it. It is certainly possible in a given case that there is an F that is not G despite the high degree of confirmation that all Fs are Gs; next year the 'trees might flourish in December', as Hume says. Nevertheless, the best science as well as common experience bespeaks the reliability of induction. Nothing higher could conceivably be claimed for it, which means – as a 'matter of definition' of the word 'rational' – that it is rational to believe it, even if we can't be absolutely certain. The principle shows why we may draw inferences about regions of the universe which are spatially remote (we've examined reality close by and infer that all of it is more or less the same in that it obeys the same laws), and regions which are temporally remote (we see how it is now and for the proximate past and infer that the remote past as well as the future are like the present in obeying the same laws).

Moreover, Ayer claims, no hypothesis can be verified on its own. Confirmation of a single hypothesis always presupposes various other conditions, such as the way that various experimental apparatuses work, or basic geometry, or the general laws of motion and so on. The French philosopher of science Pierre Duhem (1861–1916) was the first to articulate this doctrine of 'holism'; Quine would put it later as the idea that in principle it is the whole of science (or, at any rate, a large portion of it) that is at issue when testing a statement.

Second, Ayer thinks that even a statement pertaining to the sensory content of the here and now is not absolutely incorrigible, not absolutely certain. The key to understanding him here is that he distinguishes sharply between a sensation such as whiteness and a proposition describing it such as 'this is white'. He is not saying that a sensation is in any way doubtful, for that would not make any more sense than saying that a stone is doubtful. The sensation, like the stone, just is, just happens. But it can never be ruled out that the proposition, the *statement*, can mis-describe the sensation, if for no other reason than that one makes a mistake concerning what to call the sensation, in classifying it as 'white'.

So no empirical claim can be absolutely certain; they are all of them 'hypotheses'. I'll focus now on three types of ordinary statement which will show some further dimensions of Ayer's philosophy of language and illustrate its power.

Material objects. Ayer reaches back to Bishop Berkeley (1685–1753) for elements of Berkeley's *idealism*: the thesis that the world of tables and trees is composed only of ideas in the mind; that reality is mental. But he does not follow him all the way. What Ayer takes from Berkeley is **phenomenalism**: the

claim that the reality of tables and trees is the world of the *sensible experience* of tables and trees. A table, for example, is made up of sensible particulars – Ayer calls them 'sense contents'. But unlike Berkeley, Ayer thinks of sense-contents in themselves as 'neutral' between the inner world of introspection, on the one hand, and the outer world, the world of tables and trees, on the other.

The table, or rather part of the table, while I am perceiving it, is presented as a collection of sense-contents. What about when I do not perceive it (and no one else does)? Whereas Berkeley supposed that when I leave the room the table exists because there is always someone who does perceive it, namely God, Ayer upgrades the ordinary belief – that if I *did* return to my room, I *would* perceive it – to a principle governing the existence of material objects. Common sense, as well as science, tells us positively that this belief is typically true. Thus the table exists if and only if the sense-contents which constitute the table occur, or *would* occur, under certain circumstances.

Ayer calls material objects 'logical constructions', and helps himself to Russell's technique of contextual definition, of definitions 'in use', in order to explain the matter further. An *explicit* definition of 'the table' (Chapter 3, 'The theory of definite descriptions'), would be 'the table =_{df} A', for some sense-content A; but that is ruled out because a table is not itself just a sense-content: a sense-content exists only in the fleeting present and might not exist at all if no one is looking. But a *contextual* definition shows how to translate any whole sentence containing the term to some sentence not containing it. So in our case: 'the table is G iff ...', where the place of 'G' is taken by 'covered with books' or 'in my room' etc., and the ellipsis is some complicated description of the requisite *hypothetical* sense-contents – i.e. 'if I looked, then I would experience such-and-such sense contents ...'⁵

Other minds. If all one has is direct knowledge of one's sense-contents, then how can one know of the sense-contents of others? And if that is not possible, how can one know that others are conscious at all, that there are *other minds*? Ayer does not buy Mill's so-called 'argument from analogy' – 'that there is a perceptible resemblance between the behaviour of other bodies and that of my own, [and this justifies] a belief in the existence of other people whose experiences I could not conceivably observe' (129). He rejects it for a powerful reason: 'no argument can ever render probable a completely unverifiable hypothesis' – by which he means that no observation can support the hypothesis to any degree. He opts instead for *behaviourism*: other minds are (contextually) defined in terms of the behaviour of other bodies (which are defined as certain material objects, as above). The existence of material objects and other minds is verified in the same way, by the actual or hypothetical occurrence of 'an appropriate series of sense-contents' (130). Such a definition will be extremely complicated, of course, but Ayer insists that if an object behaves in every way as if it were conscious, then it is conscious, 'and this an analytical proposition', he assures us (130). To the complaint that this involves a strange bifurcation of the concept *consciousness* into first- and third-person variety, Ayer merely

shrugs it off; the only alternative to this definition is 'metaphysical'. 'The philosopher must be content to record the facts of scientific procedure. If he seeks to justify it ... he will find himself involved in spurious problems' (98).

Statements of value. These include statements of ethics, religion and aesthetics. Many ethical statements – statements central to ethical discourse and theories – are straightforwardly factual, like 'x causes the greatest happiness'. But many others – such as 'x is good' or 'x is wrong' – are irreducibly about value, and they are *non-factual*, for no conceivable observation would be relevant to determining their truth. But neither are they analytic. The function of such language is not to express proper propositions, but to *express emotions*; hence the view is called 'emotivism'. 'You acted wrongly in taking that money', for example, has only the factual content of 'You took that money'. By means of uttering the former, one is 'simply evincing [one's] moral disapproval' of the money-taking event (107). Crucial is the distinction between *expressing* one's emotion and *stating* that one has it. I may state that I'm bored without expressing my boredom, and express it without stating it by yawning. But, simultaneously, I can express boredom and say that I am bored (109), for example by saying 'I'm bored' in a languid manner. In the money case, one is *not* stating of oneself that one is in a state of moral disapproval; one expresses it. In most cases where disagreement happens, the disagreement can be settled by rational argument (not that it typically *is* so settled); one might begin with a view that shoplifting is not wrong and be persuaded to change one's mind when it is pointed out that not all shopkeepers represent the big corporations, for example. But in other cases the conflict is genuinely not amenable to resolution by an appeal to facts. In such a case there is no sense in asking which party is right; the conflict is more properly characterised as a fight, not as an argument.

The Principle of Verification is fundamental to the young Ayer; he equates it with 'radical empiricism'. Philosophy of language and logical analysis, working in concert with natural science, are the watchwords. 'Philosophy is full of questions like this', Ayer remarks, 'which seem to be factual but are not ... what is an x [or] what is the nature of x ... they are all requests for [mostly contextual] definitions' (59). 'The nature of x' is replaced by 'the use of the symbol "x"' (39). Not to see this is to be 'duped by grammar' (45). It is the same as Carnap's distinction between the material mode of speech – where one *uses* a sign – and the 'formal mode of speech' – where one *mentions* the sign. The cause of great philosophical mischief is to employ one when the other is called for.

It might have occurred to you that to identify reality with certain sense-contents makes it extremely mysterious as to how we could reckon sub-perceivable things such as quarks, or unperceivable events such as the big bang, as parts of reality. Wishing to preserve science, it seems that Ayer will have to accept some sort of 'instrumentalism' or 'fictionalism' about such things, perhaps recognising some lesser state of existence for entities that are, for their explanatory value, counted as part of our scientific picture but not as empirically real because they are not sensible. But such

views are alive and well in contemporary philosophy of science, with the most visible representative being Bas Van Fraassen.

Two other significant challenges to Ayer's view are the so-called Frege–Geach problem, and the status of the Verification Principle itself.

The Frege–Geach problem is a problem for Emotivism about value. Consider the following argument: *If shoplifting is wrong then I will eat my hat; shoplifting is wrong, therefore I will eat my hat.* One can not only understand but accept the validity of this argument, without taking a view, without bearing an attitude or feeling an emotion on whether shoplifting is wrong. But then the function of the sentence 'shoplifting is wrong' cannot be to express the emotion of moral disapproval; one's understanding of it does not depend on that. Furthermore, a valid argument form, on the orthodox view at any rate, is one such that if the premises are *true*, then the conclusion is *true*. But since only propositions can be true, then again it is not the case that 'shoplifting is wrong' merely expresses emotion. It must mean something truth-evaluable: a genuine proposition.

The challenge to the Verification Principle is simple and sweeping. Like every other statement, the Verification Principle must itself be verifiable or *a priori*. If it is verifiable, then it is possible to specify certain observations that would tend to confirm it, and others which would tend to disconfirm it. But not only does it not seem possible, it would seem circular if it were, like a snake trying to swallow itself: it looks as if to defend the Verification Principle, one would have to assume its truth. Suppose then that it is *a priori*, therefore not susceptible to verification. The trouble there is that it does not have the form of a tautology or a truth of logic such as 'No dog is a non-dog', or an analytic statement such as 'No bachelor is married': it seems to assert something substantive about the relation of language to the world. And it is a central commitment of Ayer's theory that there are no 'synthetic' *a priori* statements: no *a priori* statements which are not tautologous, logical or analytic.

● VERIFICATIONISM II: CARNAP'S LOGICAL EMPIRICISM

Ayer's work substantially overlaps Carnap's, both historically and philosophically. Carnap was older – born in 1891 to Ayer's 1910 – and whereas *Der Logische Aufbau der Welt* appeared in 1928 (translated as *The Logical Structure of the World* in 1967), Ayer's *Language Truth and Logic* appeared 1936. Carnap's *The Unity of Science* and *Logische Syntax der Sprache* appeared in 1934 (the latter translated as *The Logical Syntax of Language* in 1937), and the long paper 'Testability and Meaning' appeared in 1936 (also important for us is *Empiricism, Semantics and Ontology* of 1950; in a later chapter we'll run across his *Introduction to Semantics* of 1942 and *Meaning and Necessity* of 1947). Decidedly also the influence ran from Carnap to Ayer. Carnap was an established figure of the so-called 'Vienna Circle' when Ayer visited in 1933 (at age twenty-two) – a formidable group that involved the previously mentioned Morris Schlick, along with Herbert Feigl, Philip Frank, Hans Hahn, Carl Hempel,

Otto Neurath and Friedrich Waismann. We'll try to peel off some important differences between them to illuminate the important aspects of Carnap that were not picked up on by Ayer. We won't attempt anything like a presentation of Carnap's entire philosophy of language; we'll stick close to our theme of verificationism and ignore the many changes in his views.

If Ayer did not learn his reverence for natural science and his animus against metaphysics from the Vienna Circle and from Carnap, he had his Humean predilection for the one and against the other confirmed by his time in Vienna. In the Foreword to *Logical Syntax*, Carnap wrote:

Philosophy is to be replaced by the logic of science – that is to say, by the logical analysis of the concepts and sentences of the science of the sciences, for the logic of science is nothing other than the logical syntax of the language of science.

(Carnap 1936, p. xiiv)

He applied these anti-metaphysical ideas to none other than Heidegger, purporting to show that some of Heidegger's sentences were the worst kind of nonsense. For example, Heidegger writes:

What is to be investigated is being only and – *nothing* else; being alone and further – *nothing*; solely being, and beyond being – *nothing*. *What about this Nothing? ... Does the Nothing exist only because the Not, i.e. the Negation, exists? Or is it the other way around? Does Negation and the Not exist only because the Nothing exists? ... We assert: the Nothing is prior to the Not and the Negation. ... Where do we seek the Nothing? How do we find the Nothing ... We know the Nothing ... Anxiety reveals the Nothing.* That for which and because of which we were anxious, was 'really' nothing. Indeed: the Nothing itself – as such – was present ... *What about this Nothing? – The Nothing itself nothings.*

(quoted by Carnap, 69)

Carnap points out that Heidegger uses the term 'nothing' as if it named an entity (Ayer celebrates the point; 44–5). But anyone who knows the logic of Frege or Russell will say that 'nothing', like 'everything', is a quantifier, not a name or singular term; there just isn't any object into whose nature one might inquire. Sentences such as that which Heidegger asks about are not merely metaphysical, not merely unverifiable 'pseudo-statements', but worse: in a logically well-behaved language, no such sentences 'can even be constructed' (70).

The *Aufbau* aims to present a language in which all undefined or primitive non-logical expressions appertain directly to experiences; all sentences that do not directly report experiences are either (a) logically reducible via definitions to ones that do or (b) analytic, that is, provable from purely logical rules together with definitions (direct and contextual). Such an outline was Carnap's idea of **Phenomenological Reductionism**: the ideal, at any rate, was that every significant statement is either analytic or in principle reducible to some (possibly complex) statement

couched in the basic vocabulary of the language that directly describes experiences; hence Carnap's favoured designation for it – Logical Empiricism.

Carnap is conspicuously more rigorous and more careful than Ayer (indeed, he was one of the few who studied mathematical logic with Frege). He did original work in logic himself, and went on in later years to work on probability theory and induction. But the most interesting difference at the relevant time between Ayer and Carnap was that Ayer presents his philosophy, more or less, as the truth about language in *general*. We've seen that this runs into trouble when the Principle of Verification is applied to itself. Carnap takes another line: although he has similar views about verification, he puts forward a *formal language* governed by the principle as a 'linguistic proposal'. He thinks (for a time) that it's a fine proposal, but stops short of maintaining that all language which fails to obey it is utterly without cognitive meaning (whatever that would mean); rather, each alternative proposal can always be explored if one sets up a suitable 'metalanguage': a language for talking about the language that embodies it. Anything more would be 'metaphysical'. He thus advances his:

Principle of Tolerance. If one wishes to make theoretical claims with maximum scientific precision, one must explicitly specify the syntax of the language one is using and the rules that determine its analytic truths and inferential relations. But one cannot be criticized on theoretical grounds for one's choice of language. So even if the basic terms of the language of Carnap's *Aufbau* pertained to sensory experiences, this was not to be construed as a theoretical commitment to the idea that only such a language could serve as a basis for science.

Carnap extends this attitude towards most philosophical questions. He added two further corollaries:

Logic as convention. Neither Frege nor Russell possessed a very convincing account of our knowledge of logic itself or of the nature of logical truth. Wittgenstein – and perhaps Ayer – has an answer, but Carnap rejects it, having an uncrackable nut of ineffability built in. Carnap's answer was both simple and radical. Logical truth and the question of what follows from what are simply rules of language. To speak a language is to be bound by its rules, such as the fact that 'bachelor' is interchangeable with 'unmarried man', and similarly that an acceptance of P, and if P then Q, requires an acceptance of Q. Questions of logic are simply questions of the rules of the language one is using: all such matters are analytic, in Carnap's sense. The question of what is analytic and what isn't is a matter of convention or stipulation. One may choose whatever language one likes, depending on one's purposes. But whichever language one chooses, one thereby chooses a set of rules, and thereby chooses a logic. Furthermore, all cognitive activity – all theoretically significant thought – must be carried out within a language. Thus one is always presupposing some set of conventions, hence a logic, which, in turn, decides what trains of reasoning are valid. It follows that the choice of a language cannot be a theoretically-based one, cannot be a cognitive one. Instead, the choice is a *practical* matter. What then of Wittgenstein's claim that to explain one system of representation in

terms of another is only possible because the two share logical form which can never be explained? Carnap does not see that anything need be left out in such an explanation, for any such thing is stateable in a metalanguage; a claim that something is ineffable *in principle* is dismissed as a claim that something is necessarily unverifiable and not analytic, hence not only metaphysical but lacking in cognitive meaning, that is, nonsensical.

Internal and external. The Principle of Tolerance extends to ontology – the question of what entities a given theory is committed to. For example, we can ask existential questions such as 'Is there a prime number between 12 and 15?' or 'Is there a mulberry bush outside this window?', but we cannot meaningfully ask 'Are there numbers?' or 'Are there physical objects?' Such a question would either be what Carnap came to call an *internal* question, which can only be answered in terms of the analytic statements of the language – which are trivially affirmative and have no factual content – or an *external question*, whose answer comes in terms of the practicality of adopting the language of arithmetic or physical objects. Substantive, theoretical questions: all of them are *internal* questions. Traditional ontology – questions of what ultimately exists (numbers, physical objects and so on) – was for Carnap a pseudo-science.

● THE VIENNA CIRCLE AND THE PROTOCOL DEBATE

One tends to speak of Logical Positivism, or Logical Empiricism, not as an idea in the abstract but as a concrete historical movement. It had its peak around 1929–35, taking place primarily in Vienna, where the eponymous Vienna Circle had its meetings. Earlier figures had espoused forms of 'Positivism', such as the Frenchman Auguste Comte and the aforementioned Ernst Mach, and Russell had worked out some of the key ideas in his Logical Atomism, but the philosophical commitments of this 'neo-positivism' really came into their own with the Vienna Circle, not least because of the influence of Wittgenstein's *Tractatus* of 1921. We've discussed a bit of Carnap, and mentioned Schlick, Feigl, Frank, Hempel, Hahn, Waismann and Neurath. Ayer, as we said, visited in 1933, and in the same year Quine visited, spending some notable weeks in Prague learning from Carnap and arguing with him; we'll consider Quine's views in Chapter 6 of this book. The Vienna Circle broke up in the middle of the decade thanks to the *Anschluss* and the Nazis generally. Many including Carnap made their way to the US, thus spreading the movement further.

It was a movement in general in philosophy that was distinctly modern; indeed, as befits a movement centred on the city of Arnold Schoenberg, Gustav Klimt and Alfred Loos, it is sometimes characterised as a modernist movement in philosophy. 'The fundamental thesis of modern empiricism consists in denying the possibility of synthetic [non-analytic] *a priori* knowledge', said the Circle's manifesto of 1929. In particular it: (a) sought to dispose of 'Metaphysics' as cognitively meaningless; and (b) sought, in epistemology or the theory of knowledge, to show that science is in fact logically grounded in experience. Both goals were, of course, very much

in line with Hume's philosophy, but the new logic seemed to make it possible to actually achieve them. All agreed that a defining feature of empirical science was a certain schema: a theory is composed of theoretical statements which collectively imply certain statements of the form 'If O_1 then O_2 ', where O_1 is an observational or experimental circumstance (e.g. heat is applied to a substance) and O_2 is an observable result (e.g. the substance expands). All agreed that the experiential or observational foundation represented by O_1 and O_2 in some way concerned the senses. But they differed markedly on the details of these basic statements, on what precisely was their form and content, and on what it means to say that other statements are based upon them. We shall briefly consider three views in this so-called 'Protocol Debate': Carnap's, Neurath's and Schlick's.

Carnap. The legalistic notion of a protocol statement is an official record of an experimental observation as used in science. In *The Unity of Science* of 1934 – from an article of 1931 – Carnap was comparatively less committal as to the preferred definition of protocol statements, but in the *Aufbau* of 1928 he took these basic reports to record the whole of one sensory state – a *Gestalt* – and then, by elaborate set-theoretic means, he defined terms for sensory qualities such as 'yellow occurs' using the predicate 'recollection of similarity'. The details needn't detain us (extraordinary though they are). The main point is that the protocol statements pertain to the world of immediate sensation. The view is known as methodological solipsism – merely 'methodological' because, as we saw above, Carnap refrains from asserting that he is discovering the real primacy of the private world of immediate sensation, or that any language is the right one; it is merely a linguistic proposal. With sensory qualities defined, Carnap defines 'subjective objects' – i.e. *the-coffee-cup-for-me* – and from there, defines inter-subjective objects – public objects – such as the coffee cup (though Carnap's views changed rapidly and by late 1932 he had renounced this scheme in favour of protocol sentences as treating directly of physical objects).

Neurath. For Neurath the protocols are ordinary physical statements, but they are not simple. For example: 'Otto's protocol at 3:17 o'clock: [At 3:16 o'clock Otto said to himself: (at 3:15 o'clock there was a table in the room perceived by Otto)].' Think of the sentence within the innermost brackets as the content of a perception one had; the material in the next-most brackets '[At 3:16 ...]' expresses the corresponding verbalised thought that one had on the occasion, and the material outside all brackets 'Otto's protocol ...' is what one writes down for the official record of scientific experimental observations.

Neurath's overall motivation is that, in his view, *all* statements are corrigible. Like other statements, protocols have their conditions for acceptance. In a memorable passage, he writes:

There is no way of taking conclusively established pure protocol sentences as the starting point of the sciences. No tabula rasa exists. We are like sailors who must rebuild their ship on the open sea, never able to dismantle it in dry-dock and to reconstruct it there out of the best materials.

(Neurath 201)

That is to say, the 'fate of being discarded may befall even a protocol sentence'. For if one's theory conflicts with the protocol, and the theory is thoroughly well supported, then the protocol can be withdrawn (we've seen a similar doctrine espoused by Ayer). And to be subject to such a fate is the nature of an empirical sentence. For since protocols are not analytic they must be verifiable, and if that is not to be an empty ceremony, then they must be subject to being found false. Even the simplest language is still of the physical world; 'every language is inter-subjective', therefore is subject to the disputation of others.

Schlick. The most trenchant, if not altogether the most lucid, writing on the protocol debate was from a slightly older philosopher (who was shot dead by an ex-student and Nazi sympathiser in 1936, enraged, as legend has it, by Schlick's anti-metaphysical view, jealousy over a female student, or both). Protocols, as conceived by Neurath (and the later Carnap), cannot, in Schlick's estimation, serve as a 'firm basis for knowledge', in particular cannot serve as measuring rods for *truth*. For if they are corrigible in the way Neurath describes, then what principle do we go by in settling what our best theory is to be? It is not conformity with observation, because observations – protocols – can be withdrawn. All we can say is: the test of truth for each sentence is its *coherence* with the rest of our so-called 'knowledge'. But coherence is at least consistency, lack of contradiction. And although necessary for truth, it is hardly sufficient. Fairy stories can be told which do not harbour contradictions. Indeed, Schlick points out that *p* may be part of one internally consistent theory, *A*, and not-*p* part of another internally consistent theory *B*; if consistency were sufficient for truth, then, although I cannot consistently accept the larger theory *A*&*B*, the coherence theory will tell us that both *A* (which includes *p*) and *B* (which includes not-*p*) are true.

Schlick concludes that protocol sentences, at least in Neurath's sense, cannot serve as the foundation of knowledge. Instead, he introduces the rather Janus-faced notion of an *observation statement*. These are private, about the immediately perceived in the immediate present and 'in a certain sense cannot be written down at all' (he ought, perhaps, to have written the term in scare quotes). They are not records of mental events, but are the mental events themselves. If one seeks to confirm an hypothesis, looking for something in, say, a test tube, it is that inner moment of satisfaction, the feeling manifested by 'That's it!' 'Finality is a very fitting word to characterize the function of observation statements', Schlick writes (Schlick p. 223).

Observation statements *must* have an indexical element (Schlick calls it a 'demonstrative' element), which is the main reason they cannot be protocols. The closest thing to an example would be 'Blue here now!'; the corresponding protocol – say, 'Schlick perceives blue at time *t*' – lacks the demonstrative element corresponding to 'now', and has none of the first-person immediacy of the observation statement. It is a normal statement of ordinary language, comprehensible inter-subjectively, from the third-person point of view. Observation statements are instead private, and when the moment passes, so too does the meaning; the meaning is 'in the moment'. Thus they are like analytic statements in that one grasps their meaning and thereby grasps their truth. But on the other hand, an analytic truth affords no knowledge of the way reality happens to be; observation statements do.

● HISTORICAL NOTES

This chapter differs from the others in that the material is presented as a drama, complete with actors and even a main stage (Vienna), so historical notes would be largely redundant. But it will be useful to make two sets of remarks: one about the dangers of simplistic generalisation; the other concerning the aftermath of our story.

It's difficult to overestimate the sheer excitement and the sense of new ground being broken in what was, in retrospect, the early analytic philosophy taking place in central Europe in the post-World War I years to the mid-1930s. Not only Vienna, of course, but Berlin, the Czech Republic and Poland – as well as outside central Europe in Scandinavia, Britain, the US, France and Italy – had active, hyper-smart and scientifically minded philosophers. Many were trained in science or mathematics, rather than, or in addition to, philosophy. The general tone was anti-metaphysical, empiricist and very broadly verificationist. But it is vital to recognise that even within the Vienna Circle, there were many influences, many figures on the periphery including none other than Albert Einstein (one of the great achievements of the early-century positivists was Schlick's 'The Philosophical Significance of the Principle of Relativity' of 1915, written in close contact with Einstein). For the apparent destruction by relativity theory of the 'Kantian' picture of the mind–world relation sent philosophers of science back to their drawing boards, with conventionalism, the underdetermination of scientific theory and, of course, verificationism being explored. There were many views being expressed and many changes in the views of particular philosophers. It is common to hear of how 'Logical Positivism' represents a single view that is now dead and buried, but that is very far from being the case. If our story about Protocol Sentences contributes to that mistaken impression, then I must warn you not to think that the episode does anything like summing up the movement in general.

When the Vienna Circle and scientific philosophy generally on the continent came to an end with the rise of the Nazis and the war, most figures, as we said, emigrated: some to Britain and many to the US. Among the larger names who went ultimately to the US, which may have been familiar even before reading this chapter, are Carnap, Feigl, Gustav Bergmann, Philip Frank, Waismann and Hempel (also the philosopher of science Hans Reichenbach of Berlin). Two prodigious logicians – Kurt Gödel, of Brno in the Czech Republic but who did his doctorate at Vienna, and Alfred Tarski of Warsaw – were also closely connected with the Vienna Circle, and went to the US as well. Philosophy of Science and Logic in the US would not have had anything like its actual shape, were it not for the writing and teaching of these figures.

Wittgenstein came from a very rich and accomplished Viennese family. In 1908 as a young man, he went to England to study engineering, before switching to the foundations of mathematics and logic and meeting Russell. Russell recognised Wittgenstein's philosophical genius almost immediately – to such an extent that his own philosophy not only felt his influence, but momentarily came to a halt as Russell was so shaken by Wittgenstein's criticisms. In 1913 Wittgenstein went to Norway

to work on philosophy, then home to Vienna. He served his country bravely during World War I, and after being held as a prisoner of war after Austria's defeat, he gave away most of his inheritance to his siblings in 1919 (or rather, his remaining siblings; no fewer than three brothers had died by suicide by 1918). *The Tractatus* was published in 1921 (an English translation in 1922), and Wittgenstein immediately turned away from philosophy, leaving Vienna. In 1926, after a somewhat calamitous end to a sojourn teaching school in the country, he returned to Vienna and philosophy, where he met occasionally with members of the Vienna Circle, who were most impressed with his *Tractatus*. His main task in Vienna for a couple of years was to take on an important role in the design and construction of a house for his sister Margaret (which one can still visit as the Bulgarian embassy). Meanwhile his philosophical friends in Cambridge were encouraging him to return to England; he did so in 1929 (see the 'Historical notes' of Chapter 5).

Ayer's visit to Vienna in 1933, as we said, roughly coincided with a visit to Vienna from another youngster, W.V. Quine from America, who also went to see Carnap in Prague (where Carnap had taken a professorship). Quine was as impressed as Ayer was by the philosophy on display, but in 1936, the year of Ayer's *Language Truth and Logic*, Quine published his 'Truth by Convention', which makes a thorough-going case against Carnap and others that really there is no such thing as truth by convention, and, more narrowly, that convention cannot explain logic. Quine was as convinced as anyone that the positivists were right to disparage metaphysics – right to deny the authority of intuition, to deny that there are synthetic *a priori* truths and to deny that there is any substantive non-empirical knowledge. But he had doubts about analyticity in general, which he would develop into a still more radical type of empiricism first sketched in his famous 'Two Dogmas of Empiricism' of 1951. Some would say that Quine is the most important philosopher in the second half of the twentieth century. He was not a positivist, but he shared many of his most powerful commitments with members of the Vienna Circle and formulated some of his deepest moves in opposition to them. We'll see more in Chapter 6.

• CHAPTER SUMMARY

Russell finds fault with the view that the truth of judgements involves two abstract entities: a proposition and the fact to which it corresponds: no difference between the two can be detected. In its place, he proposes dispensing with the idea of a proposition; judgements are themselves mental facts, involving the judger and the constituents of what old scheme was the proposition, and truth is the existence of a fact involving those constituents but not the judger (unless the judgement is about the judger). Wittgenstein faults Russell's idea for being unable to explain why certain constellations of entities are not judgeable, such as the table, the penholder, the book.

Wittgenstein goes back to propositions and facts, or rather propositions *as* facts. He conceives of a proposition not as an abstract entity but as a written or spoken sentence – a 'propositional sign' – being much like a picture in that certain elements

(the names, patches of paint) as configured in the sentence show how the referents must be if the proposition is true (if the picture is accurate). Truth is explained as one fact, the depicting fact (the proposition), corresponding to another fact. The crucial difference between propositions and pictures is just that the former accomplishes abstractly via conventions (the use of predicates etc.) what the latter actually embodies (spatial arrangements going proxy for other spatial arrangements). The one aspect common to the propositional sign and the depicted fact, and to *any* representation and the depicted fact, is 'logical form'. It cannot be explained, as it is presupposed by any system of representation. This leads Wittgenstein to say that anything that cannot be pictured in this way is not literally sayable; it can only be shown, for example, in the use of language. Logical form, but also value and philosophy itself, go under this head.

Ayer puts the principle of (weak) verification at the centre of his philosophy. Meaningful propositions or statements are either *a priori* (which means for Ayer that they are mere tautologies or analytic, without empirical consequences) or there are some possible observations which would be relevant to their truth. Ayer applies the idea to several examples, including his account of physical objects, other minds, discourse about value, the past and future and idealism versus realism. The strategy in general is to find the observational evidence that we commonsensically or scientifically use to settle a matter or account for a phenomenon, and to upgrade those considerations into a criterion of the matter or definition of the phenomenon. Russell's technique of contextual definition plays a major role. The scheme has many smaller difficulties, but two major ones are the Frege–Geach problem for the emotivist theory of value and the problem that, according to the verification principle, the verification principle itself is not meaningful.

Carnap was the source of many of Ayer's ideas – especially his animus against metaphysics – but added some layers of sophistication as well as actually providing much of the formal work in logic that Ayer only describes informally. Carnap skirts the problem that the verification principle is not itself verifiable by presenting the idea only as a feature of a particular language, not as a principle governing all significant thought. Carnap's 'Principle of Tolerance' allows the setting-up of any language and the exploration of its consequences. To this he adds (1) the idea that even logical truth is a mere matter of convention; more generally, analytic truths are mere stipulations made when a language is set up; (2) the distinction between 'internal questions' and 'external questions', such as that between 'is there a stone in your shoe?' and 'are there physical objects?' The latter appears to be a philosophical existence question, but it is better viewed as a question about what language to adopt – as the question 'does this language employ terms for physical objects?' The question of what language to adopt is a practical question, not a cognitive question, for all cognitive activity must be carried out within a language. There is no language-independent or language-neutral theoretical viewpoint.

Roughly, all empirical science involves the schema 'if this theoretical statement is true, then if we contrive this observational circumstance, a certain observation will occur'. But what *exactly* are the observations? What is the canonical form of their

expression? Should the data consist of physical objects or events or items of direct sensory experience? This is a genuine dilemma: the statements have to be inter-subjective (so it seems physical, hence public), but fit to serve as a foundation for knowledge, for science (so it seems directly sensory, hence private). This was the focus of the famous 'Protocol Debate' among members of the Vienna Circle (we considered Carnap, Neurath and Schlick).

● STUDY QUESTIONS

- 1 Wittgenstein's 'Picture Theory' is an account of predication that takes propositions to be 'propositional signs' as used – i.e. actual spoken or written sentences. (a) Does the account really solve the problem of the direction of a relation (how one distinguishes 'Brad kissed Janet' from 'Janet kissed Brad')? How? (b) Does it solve the 'list problem' (the problem of explaining the 'unity of the proposition')? How? (c) Does it solve Frege's problem of the concept horse? How?
- 2 If you accept Ayer's emotivist line, then how can you explain the question 'is shoplifting wrong?' Or: suppose you were not sure whether shoplifting is wrong, and you considered the argument: *if shoplifting is wrong, then I'll eat my hat. Shoplifting is wrong. Therefore I'll eat my hat.* How can you recognise the apparent validity of the argument? One option: assume that central to logic is a general notion of *acceptance* of a (declarative) sentence, where one can accept either a factual sentence (as true) or an expressive sentence (as endorsed). So the criterion for the validity of an argument is: if the premises are accepted, then the conclusion must be accepted. Do you think that is just word-making, or is it on to something?
- 3 The Principle of Verifiability does seem reasonable, does it not? But what are those reasons? Are you giving an argument whose premises are empirical? Or are the premises you give analytic? Is the Principle subject to itself? Or does it have a special status, as a sort of definition of the 'language game' of science?
- 4 According Russell's doctrine of contextual definition, if 'the F is G' is true, then *there is* an object which is F. For example, if 'the King of France is tall' is true, then there really is a King of France. Contextual definition does *not* allow one to speak as if the object exists when really it does not. But Ayer seems to say that logical constructions allow true speech using 'the F' (such as 'the table') when strictly speaking there is no object which is F (a table); there are just certain sense-experiences that would be had if one were in such-and-such a situation. What are these 'hypothetical sense-contents'? Are they *possible* contents? Or does Ayer require a more radical style of definition, so that the defined objects are really 'logical fictions'?
- 5 Is Carnap merely trying to disown his responsibilities in saying that the Verification Principle is just a proposal?
- 6 Who comes off better in the Protocol Sentence debate, Neurath or Schlick? Schlick does have a point against Neurath, no? However, some twenty years later, the American Wilfrid Sellars launched an attack on 'the given', which Schlick looks to be committed to (more recently, John McDowell joins with Sellars). In his quest for the true foundation of knowledge, Schlick is forced to say that what

he called 'observation statements' cannot be expressed in our public language; these are private experiences. But how then can these experiences engage with ordinary sentences of public language? How can they bear rationally upon them? The relation, it seems, must be logical – but in that case, they are *not* sealed off from the reach of language. How do you see the issue between Schlick and Neurath (and perhaps Sellars)?

● PRIMARY READING

Ayer, A. (1946) *Language, Truth and Logic*.

For an introduction to various members of the Vienna Circle, see the collection by A. J. Ayer, editor (1959) *Logical Positivism*.

Carnap, R. (1928) *The Logical Structure of the World and Pseudoproblems in Philosophy*.

Carnap, R. (1956 [1947]). *Meaning and Necessity: a Study in Semantics and Modal Logic*. This includes the paper *Empiricism, Semantics, Ontology*.

Russell, B. (1910) 'Knowledge by Acquaintance and Knowledge by Description,' *Proceedings of the Aristotelian Society*, 11: 108–128. Reprinted in his (1918) *Mysticism and Logic*, pp. 152–167.

Wittgenstein, L. (1961) *Tractatus Logico-Philosophicus*.

● SECONDARY READING

Coffa, A. (1991) *The Semantic Tradition from Kant to Carnap: to the Vienna Station*. A very readable account.

Richardson, A., and T. Uebel editors (2007) *The Cambridge Companion to Logical Empiricism*. Authoritative compendium with top-drawer contributors.

Uebel, T. (2007) *Empiricism at the Crossroads. The Vienna Circle's Protocol Sentence Debate*. Details of a fascinating story.

● NOTES

1 Axiom V: The set of F's = the set of G's iff: For every x, x is F iff x is G.

2 This is not to say that the idea of truth-as-correspondence-to-fact cannot be made to work. A good book that develops the idea is Barwise and Etchemendy (1987).

3 Wittgenstein call elementary facts 'states of affairs' (German: 'Sachverhalt'), reserving 'facts' for compound facts (German: 'Tatsache').

4 Wittgenstein's actual account involves just a single truth-function ' $N(\dots)$ ', which takes any number of propositions as arguments; ' $N(p_1 \dots p_n)$ ' is true iff all the arguments $p_1 \dots p_n$ are false. 'Joint Denial' is truth-functionally adequate: any truth-function – including conjunction, negation etc. – can be defined in terms of it.

5 See the study question number 4 if you spot a problem with this.

5

• the late Wittgenstein

Ludwig Wittgenstein wrote two books: the *Tractatus Logico-Philosophicus* (1921) and the *Philosophical Investigations* (published posthumously in 1953). Superficially the two are similar: each contains a philosophy of language, and neither is organised into continuous prose which carries the reader along an argument to a conclusion. Each is an assemblage of remarks, seemingly arranged by cutting and pasting, as if the rich variety of connections between remarks resisted the finality of linear ordering.

Still the difference of substance between the two could not be greater – at least, so it has seemed to many readers. For our purposes, we can think of the younger Wittgenstein, the author of the *Tractatus*, as accepting the essentials of the Frege–Russell view of the functioning of language. As we saw in Chapter 4, the picture theory of the atomic sentence found in the *Tractatus* retains the Frege–Russell view that reference is fundamental to language, to objectivity and to the relation between the mind and the world – even though in the *Tractatus*, Wittgenstein disagrees with both Frege’s view and Russell’s in seemingly profound ways. However, from the point of view of the later Wittgenstein, the author of the *Investigations*, those disagreements can seem inconsequential. He no longer thinks of reference as anything like a linchpin for understanding language, and indeed thinks there is no systematic and definitive way of mapping the multiplicity of language, in all its unfathomable diversity. To look for systematic generalities would be like imposing a grid onto a medieval Italian city, with all its twisting lanes, dead ends, irregular *piazzi*, tunnels, bridges, even streets atop streets. You could do it, but it would be arbitrary and distorting, and you would not thereby get to know the city in the intimate way you would by exploring it on foot. And it is this later, much more radical view that this chapter is about.

• LANGUAGE GAMES

According to Frege, Russell and, indeed, most philosophers of language, including the younger Wittgenstein, reference or naming is fundamental. In the words of the

Tractatus, that is the point at which language reaches out to reality, where it touches the world. Other words that play fundamental roles in the classical conception are ‘understanding’ and ‘meaning’. According to the later Wittgenstein, these words foster mere illusions of clarity, as when we say, as if we were reciting a religious slogan, that to *understand* a sentence is to *grasp* its *meaning*. On the contrary, looked at from ‘close to’, understanding proves to be a multiple-dimensional business (what counts as understanding one word may be quite different from understanding another), and there is no way fruitfully to cut out the specifically linguistic capacity for dissection and investigation from other skills and competencies. There is no one thing that constitutes understanding – it depends on details of the case – and understanding something is not all-or-nothing but a matter of degree.

This emerges clearly when we consider one of the **language games** that Wittgenstein describes at the beginning of the *Investigations*.

I send someone shopping. I give him a slip marked ‘five red apples’. He takes the slip to the shopkeeper, who opens the drawer marked ‘apples’; then he looks up the word ‘red’ in a table and finds a colour sample opposite it; then he says the series of cardinal numbers – I assume that he knows them by heart – up to the word ‘five’ and for each number he takes an apple of the same colour as the sample out of the drawer. – It is in this and similar ways that one operates with words.

(1953, §1)

Although of a humdrum sort, the skills that the shopkeeper and the shopper have to master in order to play such a game are inescapably diverse, and it seems somewhat arbitrary to call some of the skills *linguistic* skills and others not. There is no hard and fast division. Taking the list to a certain place and handing it over, finding the drawer marked ‘apples’, attending to a colour sample and responding appropriately, taking the correct apples out of the drawer and so on – these constitute what Wittgenstein calls ‘**forms of life**’. They are partly linguistic, partly not, and have the full significance they have through their being embedded in a larger framework of overall life and culture. To isolate and pull out linguistic phenomena for examination is to look at parts out of context, as if we were examining a disembodied gallbladder on a lab bench, trying to describe it while remaining completely ignorant of the role it plays in the body.

The point of calling such activities ‘language games’ is not to disparage the activities or to take them lightly. But part of the point of calling them ‘language games’ is to stress the affinity between the words employed and pieces in a game. The significance of identifying a certain piece as a *rook* depends upon a certain framework, namely the overall rules and practices of the game of chess – which collectively describe the moves that one may make with a rook. Furthermore, in playing chess one doesn’t just follow the rules. Within the rules there are unendingly many opportunities for creativity and drama, and likewise in the case of language games, except the overall framework is *much* wider and less definitely bounded than criteria for reference or meaning. Insofar as the rules lay down necessary conditions of saying

something intelligible, they are often treated in practice as themes for variation, as means that can be altered for some end.

This is not itself to criticise normal linguistic description. One could correctly *say* ‘five’ refers to five, ‘red’ to red and ‘apples’ to apples, and normally we’d understand the point of one’s saying these things (if there is such a point). But one would not be adding anything to the above description of the language game; one would not thereby be explaining what is really going on. On the contrary, to suppose that one would thereby be identifying what is really going on is to submit to the *craving for generality* embodied in such words as ‘refers’, ‘signifies’ and ‘means’. Such words do serve the purpose of summing up various linguistic activities, but the insight so afforded is only of a superficial sort, precisely because it is so general. The job of such a word is as much to conceal differences as to bring them to light in a way as to afford insight (‘I’ll teach you differences’, said Wittgenstein, quoting a line in *King Lear*).

The proper role of such expressions as ‘refers’, ‘signifies’ or ‘means that’ is to indicate, *once a background is understood*, or once a place has been conceptually carved out, what particular word one is to use for a certain purpose. A somewhat artificial example, but one that makes the point clearly, is the case of a starter’s saying ‘Set!’ to a field of sprinters about to commence a race: it is clear that it is only against a complicated background of rules, expectations and traditions that that syllable, uttered in that exact context, has the significance that it has. Likewise if we say of the language game described above that ‘red’ refers to redness, it is only because so much has been presupposed, only because we see the relevant background. Normally, we share so many forms of life that we never notice them; they go without saying. Compare walking down a busy street while carrying your shopping and talking on your smartphone: it is in fact a very complicated business, but we cope effortlessly because we are so utterly used to it. Again, the semantical vocabulary has its point only when forms of life, comprising various habits and dispositions, are understood as being in the background.

Another important thing to notice about the language game is that, as described, it is *complete* – the point is not that more detail could not be added to the description, but that one could perfectly well imagine such a skill set being the only one mastered or rather the only one mastered that one is tempted to call linguistic. There is no reason to maintain that the language game is incomplete or partial. Shopkeepers fully understand the words they are using; they are fully competent in the activities carried out by means of them, and that is that.

From this point of view, the somewhat humorously disparaging attitude Wittgenstein takes towards Russell’s and his own earlier idea of philosophical analysis, as culminating in logical simples, makes perfect sense. In §60 of the *Investigations* we read:

Suppose that, instead of saying ‘Bring me the broom’, you said ‘Bring me the broomstick and brush which is fitted on to it.’!—Isn’t the answer: ‘Do you want the broom? Why do you put it so oddly?’—Is he going to understand the further analysed sentence better?

(1953, §60)

● FAMILY RESEMBLANCE, TOOLS AND CITIES

Suppose a child has mastered the shopper's role in the above language game. If we now imagine the child as mastering more and more language games – each perhaps borrowing some elements from others but also requiring the acquisition of new forms of lingo – we have a picture of the child's overall skill set as consisting of various cells, each interwoven or related to others but having its distinctive territory, its distinctive cluster of purposes. The language of storytelling and make-believe, of being a pupil at school, of joking, of professing wants, of making plans with a friend, of asking favours and so on: there is no one place at which we declare the child a master of language. It's a matter of degree. And learning language involves many capacities, many kinds of mastery, as just intimated. What, then, *is* language? The answer is that there is no one thing. Wittgenstein is 'anti-essentialist' about language. No neat definition is possible. Instead, the concept of *language* is a **family resemblance** concept.

There are other examples of family resemblance concepts. Consider the concept of a *game* (Wittgenstein's celebrated example; and, of course, there is a kind of double irony in view of his use of the idea of a 'language *game*'). What exactly is a game? What are the necessary and sufficient conditions for something being a game? According to Wittgenstein, the question has no answer. Yet the concept, or rather the word, is none the worse for it. Some games involve a board and pieces (backgammon, Parcheesi), others don't; some involve competition, others don't (solitaire, charades); some have an element of luck, others don't (chess, noughts and crosses or tic-tac-toe); some are for the diversion of the participants, others not (professional sports). Instead, there are various features that are sufficient for being an instance of the concept: things having attributes A and B, or B and C, or C and D – but there are no features distinctive of or common to all games. Nor is it right to say that the concept is *disjunctive* in character, as if we could define it thus: something is a game if and only if it has attributes A and B, or B and C, or C and D, or ... for some finite number of terms, covering all actual games present and past. Besides being a case of merely 'playing with words', as Wittgenstein says – and note that no one is likely actually to come up with such a disjunction – this misses a fundamental feature of family resemblance concepts, namely that the concept of a game is *open-textured*. Even if we had such a definition involving an exhaustive disjunction as above, such a definition fails to reckon with past possibilities that went unrealised, and with *future* possibilities. If the linguistic community is faced with a new thing – with a novel set of attributes which does not quite qualify it for inclusion under the concept as it has been applied in the past – the concept may nonetheless be stretched to the new thing, if the community sees fit. Thus with time the concept may expand its field of correct application; the model thus provides for conceptual change over time, for the dynamism of linguistic expressions.

Once one has the idea of a family resemblance concept, one can seem to spot them all over the place, often in philosophically sensitive areas – potential examples include *goodness*, *knowledge* and *art*.

Returning to the concept of language, ‘language’, as we said, itself expresses a family resemblance concept. No one thing is common to the sundry activities we characterise by means of the word. Instead of thinking of language as something of which the essentials can in principle be characterised at one fell swoop, Wittgenstein suggests an analogy with cities – not planned ones, such as Milton Keynes or Brasilia, but organic and historical ones, such as London or Mumbai. The former are planned out in advance of being constructed and used, sometimes by a single person (a good analogy for these would be the notation of chemistry, or logical calculi). The latter, just like a natural language, evolved bit by bit, over time (and it is not necessary for any single person to know all of it). There need not be an overall logic to its development. London developed over a long time (over roughly 2000 years), higgledy-piggledy, sometimes expanding rapidly, sometimes undergoing periods of stasis, other times shrinking – likewise the English language. Both are organic, natural products of our evolution. There is no overall grid to be abstracted out. Still, you can get from one place to another: the streets are not all dead ends (English grammar is not hopeless). And certain parts – normally the suburbs but sometimes the centre, as when Haussmann redesigned central Paris – are self-consciously designed; in our language, the scientific bits are most conspicuously like that.

Thus we can see various language games as woven together in diverse ways and a child acquiring language as finding his feet with each of them, some earlier than others, some presupposing others, some having to be learned simultaneously with others; it is a piecemeal business. Thus, as you might expect, Wittgenstein does not regard what we have been calling ‘cognitive language’ – fact-stating language, the stuff of assertion, truth-conditions and logical implication – as taking peculiar precedence in the order of what in the Introduction we called speech-acts (and which is a topic to be returned to in Chapter 9). In language, in addition to assertion or statement, we have greetings, insults, instructions, commands, betting, haggling, moral praise and censure, aesthetic praise and criticism, poetry, banter, warnings, jokes, slang, simile, metaphor, threats, exclamations, dirty words, words we cry out under duress, in fear, in disgust, in pain, when victorious, in encouragement ... and many other things. Expressions exist for all these purposes, and some are not *referring* terms at all – ‘Goodbye’, for instance, or ‘Come on!’ or ‘It’s raining’. Again, it would be very hard to say what they all have in common; certainly, we cannot assume *a priori* that they do have something in common that is of theoretical importance.

The variety of language games Wittgenstein points to makes it unsurprising that he comes out with the famous saying that: ‘For a *large* class of cases – though not for all – in which we employ the word “meaning” it can be defined thus: the meaning of a word is its use in the language’ (1953, §43). Rather than impose a specious uniformity on language by saying such things as that every word expresses a *meaning*, it is often more instructive and less misleading to ask after the diverse *uses* of words.

Wittgenstein famously compared language to a toolbox, and words to tools. If we tried to say what all tools have in common, we should end up with something totally uninformative and probably false, like ‘all tools are used to alter something’ (or even: ‘A tool is a thing that is *used*’). Think of the hammer, saw, measuring tape, vice,

screwdriver, glue pot, paintbrush, pliers, plumb bob, socket wrenches, soldering iron, sandpaper ... and so on. Philosophers often talk as if the stating of facts or the conveying of factual information – and related acts such as questions, understood simply as requests for information – were the sole purpose of language. That would be like someone who looked into a tool box and saw only the hammer and the nails. There are certainly limits to this analogy; or, at least, it can be question-begging to push it too far. Most conspicuously, we should not assume that all the purposes served by language could be conceived or explained independently of language. We can always explain the intended result of using a tool in a way that makes no reference to tools (except when using tools to make or repair tools). But it might well be, for example, that our ability to think some kinds of thought depends on our mastery of language; so when we use language to record or convey certain sorts of thought, the intended result – perhaps that a certain person should acquire a certain belief – cannot be explained except in terms of that person's mastery of language. Perhaps many linguistic activities are themselves inextricable from Wittgenstein's forms of life: those are convention-bound practices that make up certain departments of living and without which certain kinds of purpose and value could not exist.

● TO FOLLOW A RULE I

We now come to a Wittgensteinian topic that has received a mountain of commentary, especially since Kripke's *Wittgenstein on Rules and Private Language* of 1982. The topic has become frighteningly complex, but the essentials are quite simple.

We'll take an elementary example, but it will serve to make a point that is of completely general significance. Suppose one is walking through a corridor at an airport, and one comes to a 'Y' junction. There appears the sign:



Of course, without any hesitation, you go right. The sign, as we say, *means that* one should go right; that's what it tells us. But how do you know? What fact do you know? What does your understanding consist of?

One might point to a mental occurrence; perhaps a mental image occurs when you understand. But such a mental object or state is not necessary or sufficient for being competent with the sign. It is not necessary because one can perfectly well imagine one's being competent with the sign without an accompanying mental image occurring; one sees the arrow and one goes right, but there is no need to suppose that alongside these events is another mental process that explains the event.

But, more importantly, it is not sufficient for competence; the existence of such a mental occurrence cannot explain one's understanding. For if something, so to speak, were to flash on one's inner mental screen at an appropriate moment, that something would be in exactly the same situation as the original sign on the wall. It would have to be read, interpreted; the question of *its* meaning would arise.

Suppose, for simplicity, that an internal, mental duplicate of the external, physical sign was thought sufficient for understanding the latter. Then the question of the interpretation of the inner sign arises. It could mean that one is to go *left* – a society in which arrows point the opposite way from the way we take them to point is perfectly imaginable; as is a society in which that particular mark meant ‘do not enter’ or ‘do not use mobile phones (cell phones)’. It’s a convention. So merely having another sign appear on one’s mental screen is insufficient for understanding the outer sign. But so it is with anything that might happen in one’s mind, no matter how complex in comparison with an inner replica; such inner phenomena cannot constitute understanding. ‘No [mental] process’, writes Wittgenstein, ‘could have the consequences of meaning’ (1953, II: 218).

Neither Wittgenstein’s reaction nor his reason is scepticism: he doesn’t suggest that maybe no one understands anything or that the whole idea of understanding and meaning is a fraud. Instead, he says that we must grasp signs in a way that is not an interpretation at all. Nor does he mean that understanding is magic. If ‘interpretation’ is understood as seeing one sign in terms of another, then this follows straightforwardly from the discussion of the last paragraph. Rather, learning to follow or obey a sign is a matter of *training*, or rather of acquiring dispositions such as to go right in such circumstances as encountering the sign in the airport. ‘Explanations have to stop somewhere’, he points out early in the *Investigations*; digging down, one reaches bedrock; one’s ‘spade is turned’. Responding as we do is just what one is ‘inclined to *do*’; it is just how we roll. And the capacity to learn in this way is a feature of our natural history. Other sorts of creature, with very different natural dispositions, would simply be unable to respond to the training as we do. The possibility of communication depends on ‘agreement in judgements’, as Wittgenstein puts it (1953, §242): if we were not already disposed to respond in similar ways to events, communication could not get started. The comparison is often made with the slave boy in Plato’s *Meno*: by teaching the slave boy a proof by merely asking him a sequence of questions, Socrates purports to demonstrate that knowledge of geometry is innate. Maybe Wittgenstein would not agree that *knowledge* of geometry is innate, but maybe he would agree that Socrates does show that certain pre-existing dispositions are necessary (for learning geometry).

● TO FOLLOW A RULE II

Does having a disposition to respond to a sign constitute understanding the sign? No; at the minimum, one has to be disposed to respond to it *in the right way*. At the very least, one has to be disposed to respond to it in ways that other members of the linguistic community do. This raises an issue that forms the heart of Saul Kripke’s book *Wittgenstein on Rules and Private Language*, of which everyone studying Wittgenstein should have some inkling. We can separate it into two parts.

First, a person has necessarily been exposed to a given rule on only so many occasions, in comparison with the infinite range of uses of the rule, actual and potential. One can never know with absolute certainty that one has the right disposition;

perhaps the disposition one has, and the use actually indicated by the rule, have matched in the occasions that one has so far encountered, but branch apart in other cases that one has not so far encountered. So there is a sceptical problem, an epistemic gap; how can one ever know that one has got the rule right?

Second, there is a **normative** gap between having a disposition to *respond* to a sign as others do and *understanding* the sign. Even if I am disposed to go on in what is in fact the right or correct way, it does not follow that my doing so is itself right or correct. For dispositions, in themselves, have no normative force. That a creature behaves with a certain disposition doesn't mean that they are *right* to do what they do; likewise having dispositions cannot tell me what I *ought* to do, what is *correct* to do, any more than a dog is doing the right or correct thing when he moves his leg up and down when you scratch his back. Nor is being able to predict what one is going to do. Correctness is the whole point of speaking of *rules*, of meaning. On the one hand, I might make, systematically, mistakes; that is a case where my performance is systematically incorrect according to the rule, which should not be possible if having a disposition were all that was involved in obeying a rule (for, trivially, one is disposed to act as one does act). There is a difference between a person who grasps a rule but makes mistakes and a person whose behaviour is exactly the same but isn't following any rule. On the other hand, even if my behaviour accords exactly with the rule, it might be a case of *merely* acting so as to match the rule rather than being *guided* by it (a waffle iron, even if it performs perfectly, is not following rules).

How Wittgenstein responds or would respond to these points has been the subject of much controversy. Kripke's own solution, presented on behalf of Wittgenstein, has won a fair following, but many others have been proposed. I will just sketch what I personally believe was more or less Wittgenstein's own response to the second difficulty. The key is the social context in which such linguistic behaviour takes place. Among the various language games is the habit of correcting people when their linguistic dispositions run off the rails. It is this element of our practice that accounts for normativity, and nothing more; we admonish and criticise such people (or even punish them). Without such a community it would not make sense to speak of speech as being right or wrong – even if it seemed to the language-user as if it did. Yes, this does mean that just as the individual cannot have dealt with all questions, actual and potential, of interpretation of a given sign, neither can the whole community. They can have dealt with a great deal more such cases but not all. But the challenge was only to allow for normativity in the community's practice, which the present suggestion does. As for the first difficulty, I do not think that Wittgenstein would be disturbed by it. He is no more bothered by such things than he is about the principle of induction, the problem of justifying normal inferences from observed cases to unobserved cases. Just as one cannot prove that the sun will tomorrow rise in the east, one cannot prove that our linguistic forms of life will go on as we expect. Beyond that, the issue is a worry over something that can never be realised: the survey at a given time of all possible cases. He recommends not a solution to 'the problem' but quietism, that this is a place where we must simply acquiesce in forms of life. But that is a long story.

● PRIVATE LANGUAGE

At §243 of the *Investigations*, Wittgenstein introduces the idea of a language in which words stand for inner experiences, things that can *only* be known to the person speaking (or the person writing down the words). He then proceeds to cast doubt on the intelligibility of this – a **private language**, a language that is necessarily understood only by the speaker. He does not mean to disturb the evident fact that we use words to talk about our inner experiences so long as the words can, in principle, be understood by others (more on that in a moment). He means to argue against the idea of a language that is *necessarily* private.

Suppose you try to construct such a language. You have a sensation which you call 'S'. You note it down, intending that 'S' labels S. Later a sensation intrudes. Is it S? One's natural response is to say that one can simply compare it inwardly with the previous case. The problem, however, is that such an inner comparison makes no sense. There is no **criterion** available to answer the question. For there is no gap between the *existence* of an inner sensation and its only *seeming* to exist. One cannot say: if it *seems* to be S, it's S; if not, not. For a word to be meaningful, it has to be conceivable that one could *mistakenly* misapply the word. It has to be conceivable that it seems to be S but is not in fact S. If there is no such thing, then one's attempt to mean something by 'S' has failed.

As Wittgenstein writes:

in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'.

(1953, §258)

And:

'obeying a rule' is a practice. And to *think* one is obeying a rule is not to obey a rule. Hence it is not possible to obey a rule 'privately'; otherwise thinking one was obeying a rule would be the same thing as obeying it.

(1953, §202)

We can ask: what is the criterion for the sameness of a sensation? What is the criterion for re-identifying S? Or: what is the criterion for being S? Without one, no standard for 'being S' is determined, so no meaning. Thus the argument, laid out succinctly is:

- 1 Meaning something by a word requires a distinction between applying the word correctly and only seeming to apply it correctly.
- 2 A private word necessarily lacks that distinction.

So: no private word is meaningful.

When *Investigations* first came out, many readers supposed that the impossibility of private language implied that sensations are necessarily *public* – or that Wittgenstein was arguing for a crude sort of behaviourism. However, the matter is more subtle than that. Learning, for example, a pain-word is indeed learning a bit of behaviour, but it is acquired in connection with pain and its accompaniments:

words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour.

(1953, §244)

Again, words have their meaning in context. They are interwoven in the rich matrix of life and have their roles only within the whole. We said earlier that it is the fact that we have pre-existing dispositions that makes a language possible. Included in these are ‘natural expressions’, as Wittgenstein describes; words for inner sensations, such as pain, ride atop them. Wittgenstein is not behaviourist in the sense of one who denies that inner states exist; it is that the *criteria* of words for inner states must pertain to the behaviour of human beings. This should not be surprising: at the very beginning of the *Investigations*, Wittgenstein stresses that the ways of language are explicable in terms of public practices, forms of life.

We can connect this with a doctrine previously encountered. Russell’s theory of language, in the form we studied in Chapter 2, rested on his notion of sense-data and his principle of acquaintance. ‘Every proposition we can understand is composed entirely of elements with which we can be acquainted’, runs the principle of acquaintance. Sense-data are the things with which we are ultimately acquainted, and they are the means we have of referring to physical objects. Sense-data are private entities *par excellence*. Clearly Wittgenstein raises severe doubts about such doctrines – including perhaps the more modern notion of *qualia* in the philosophy of mind.

● HISTORICAL NOTES

(See the ‘Historical notes’ to Chapter 4 for the first half of Wittgenstein’s career.) Soon after Wittgenstein returned to philosophy in 1929, he returned to Cambridge and would soon develop the radically new way of thinking displayed in the unofficial manuscripts known as ‘The Blue Book’ and ‘The Brown Book’, circulated in the 1930s and which would eventually issue in the *Philosophical Investigations* in 1953, two years after Wittgenstein’s death. The number of books on Wittgenstein’s philosophy is enormous. Many of the most penetrating are written from the point of view of disciples, who treat Wittgenstein’s writings almost as if they were sacred texts. That is by no means to belittle his disciples: it is just testimony to the quality of and fascination held by the writings. Even among philosophers who have not been quite so keen to follow explicitly in his footsteps, there is a wide range of figures – from Crispin Wright to John McDowell to Stanley Cavell – whose work has been profoundly shaped by Wittgenstein.

● CHAPTER SUMMARY

Wittgenstein proposes a quite different way of thinking about language from Frege's, Russell's or that of his own former self. Language, according to the later Wittgenstein, may be thought of as a fabric made up of language games, without presupposing that it contains a single unifying thread. *Language* is a family resemblance concept, not a concept such as 'bachelor', which admits of definition in terms of necessary and sufficient conditions; languages are organic entities, protean and diverse, with parts criss-crossing and overlapping. The concepts of reference, meaning and the like distort our understanding of these games, if we try to use them to theorise about all such activities in abstraction. If we 'look and see', for example, at the simple example of the shopper and shopkeeper, we find that such language games can be described without those concepts being privileged in the description. Their applicability presupposes a great deal of stage-setting, but once the stage is set, they do not contribute nearly so much as one might suppose.

A great illusion is that the meaning of a word is queer and mysterious, like a ghost attached to the word, telling us how to respond to it. Careful attention to cases reveals that such a thing would be powerless: if, for example, the meaning were a mental entity that was grasped whenever one used or heard the word, the entity would stand in need of interpretation, just as the word itself does. There is no choice but to accept that understanding is a matter of grasping *without* interpretation; at the bottom-level case, one is *trained* to respond as one does – one catches on, just as in learning to tie your shoes. Such training results not in mere behaviour, like a trained parrot, but in understanding, because one learns to respond against the background of a linguistic community, whose ways constitute the norms, the standards of correctness, that one strives to match.

Russell along with many others said that the ability to talk about material objects depends upon our acquaintance with sense-data. Wittgenstein argues that such a language is impossible. For meaning something by a word requires a distinction between applying the word correctly and only seeming to apply it correctly; a private word – for example, one for a sense-datum – lacks that distinction.

● STUDY QUESTIONS

- 1 Looking back at the start of this chapter, a good question to ask is how or whether Wittgenstein could cope with the demand represented by the principle of compositionality. For in dis-enshrining meaning he appears to remove the one semantical concept in terms of which the compositionality challenge was to be met. How can one respond?
- 2 Think of language as used on a factory floor, as used among lovers or intimate friends, as used on the street or in the pub, as used in the court of law and on various other occasions; surely there are norms at work – perhaps Gricean rules – but do they change, depending on context? How? In what manner are they understood by the participants?

- 3 Does Wittgenstein's picture overplay the differences, and underplay the similarities, among different types of language?
- 4 John Stuart Mill proposed the following solution to the problem of other minds (how does one know that they exist): when I'm in pain I cry out. Other human beings are very similar to me. Therefore when others cry out, I can infer by analogy that they are in pain. Having read a bit about the private-language argument, what do you think of Mill's solution?

● PRIMARY READING

Wittgenstein, L. (2009) *Philosophical Investigations*, fourth edition.

● SECONDARY READING

Ahmed, A. (2010) *Wittgenstein's Philosophical Investigations*.

Lugg, A. (2000) *Wittgenstein's Investigations 1–133: A guide and interpretation*.

6

• quine's philosophy of language

Wittgenstein's is perhaps the most famous name in analytical philosophy of language for those with only minimal acquaintance with the field. Another big name in analytical philosophy of language – historically but also conceptually – but which is much less well-known outside the field, is that of W. V. Quine. Despite his enormous influence on the philosophy of language since World War II, his actual theory is something of an outlier. And a great deal of what he is most famous for is seen by many as unjustifiably critical or negative, and felt as an irritant that needn't stand in the way of the positive or constructive proposals they advance. Certainly by 1980, the most popular areas of the analytical philosophy of language had shifted to views diametrically opposed to Quine's view. But most Quine scholars agree that these developments fail to acknowledge the full force of his position, fail to take stock of how the different aspects of his philosophy reinforce one another. We'll try to gain some appreciation of both the destructive and constructive sides of Quine, especially of how those aspects support one another.

• QUINE'S NATURALISM

Quine's interest in language comes via an interest in knowledge. We have physics, chemistry and mathematics as exemplars; what exactly is involved in knowing them, and what exactly are their existential commitments? These are basic questions of epistemology and ontology. Very much in keeping with Frege's and Russell's approaches, Quine approaches the questions via language: language is the medium of knowledge, so we want to know more about what is involved in our having it, especially the aspects most directly relevant for representing it.

From a certain point of view, however, such philosophies of language as Frege's or Russell's can seem rather high-falutin'. They seem to make substantive *a priori* assumptions

about language, its relationship to the world and its relationship to the mind. These include assumptions about what sorts of concept – such as *sense*, *understanding* – could be assumed as sufficiently clear, and as determinately applicable to actual phenomena, to utterances in the real world. Quine's main project in the philosophy of language was to avoid making those sorts of assumptions. Instead, we should examine what happens if we try to approach language, and especially the theory of meaning, from the point of view of natural science. Physics does not use the ordinary concepts used to describe the physical world in formulating itself as a rigorous science. What would a rigorous science of language be like? Can the concept of meaning be made scientifically respectable? Frege and Russell spoke of meaning and of meanings, of grasping propositions, of reference and truth-conditions, as if the existence of those things, and the explanatory value of those concepts, could be either presupposed, or established *a priori* by philosophical reflection. But they did not say very much about how such concepts can be known to apply to objects and events in the real world; they just assumed that they have determinate application or stipulated their application in laying out their formal or symbolic languages. Yet language and its meaningfulness, surely, are phenomena to be met with in the real world of noise-making animals; shouldn't a theory of meaning describe what makes it the case that a certain noise or mark has the significance that it does? Do we really have to assume that immaterial meanings or propositions explain human cognition, the stuff of axons and synapses?

Quine asked that sort of question, and his answer, to cut a long story short, is that there cannot really be a science of meaning of the traditional sort, which assigns particular meanings to words: practically speaking, the notion of meaning is probably indispensable – that is, we need it to get by in our dealings with each other – but from a rigorous, scientific point of view there is really no such thing. Intuitive ascriptions of meaning are inescapably pragmatic and interest-relative, and none are uniquely justified by any conceivable objective evidence. How we describe the meaning of a statement is, to a significant degree, a matter of convenience, of our always shifting interests, and there is no theoretically useful way to make more of it than that.

Quine promotes a general philosophical point of view which he terms **naturalism**: it is the sum of natural science, and not philosophy conceived as an *a priori* discipline above or behind science, that sets the standard for knowledge and reality. Philosophy from this point of view is rather a hodgepodge without a defining essence; but in general it has its legitimate place by virtue of its generality, its abstractness, and especially its penchant for those problems, paradoxes and confusions thrown up by the human endeavour to acquire knowledge, which for some reason or other continue to resist clarification. There is, Quine famously proclaims, no 'first philosophy'; it is only within science itself, and not in some prior philosophy, that reality is to be identified and described' (1981, p. 21). There is no rock-solid foundation for knowledge, of the kind traditionally assumed. Quine often uses an image from the Austrian philosopher of science Otto Neurath (1882–1945): as scientists or philosophers we all are aboard a ship at sea, having to make running repairs, but it is only a philosopher's conceit to think we can deconstruct the whole ship without sinking (see Chapter 4, pp. 86–87). There is no knowledge that is somehow unscientific in principle. Scientific knowledge is not different in kind from ordinary knowledge. If some discipline

resists being subject to the norms and standards prevailing elsewhere in natural science – caricature examples include astrology and homeopathy – then that is a reason to doubt whether it's really knowledge.

As we said, however, Quine's outlook is not merely negative. He is very much concerned to propose a positive conception of language, one that can account for communication, especially those aspects of language that house our beliefs and knowledge, in ways that are immune to the sorts of objection just mentioned.

● THE FIELD LINGUIST

In something like Wittgenstein's way, Quine reasons that language, and in particular meaning, must be public. 'Language is a social art', he declares (Quine 1960, ix). He disparages a 'mentalist' theory of meaning, an account which consigns the crucial facts of meaning to a private mental realm. 'Language, we are told, serves to convey ideas', he writes; 'Now how do we know that these ideas are the same? And, so far as communication is concerned, who cares? We have all learned to apply the word "red" to blood, tomatoes, ripe apples, and boiled lobsters. The associated idea, the associated sensation, is as may be' (Quine 1974, p. 35). However, even reference – the application of the word 'red' to blood etc. – cannot directly be perceived. We can only observe simple sentences like 'That is red!', said in the presence of blood. In order to generate a theoretical account of reference, then, we shall at least want a procedure which, for any speaker, describes the speaker's use of declarative sentences, which can in principle account for the speaker's referential language.

Central to this is Quine's notion of a linguistic disposition. Suppose you had a register that described *all* of the linguistic dispositions of a given human being, at a certain moment. Call the person Jo and the moment *t*. The register needn't predict what sentences if any Jo would actually say at moments at *t* or directly after *t* – that depends partly on what is happening in and to Jo at *t* – but it would tell you for each sentence of Jo's language and set of circumstances whether Jo *would* assent to it if asked under those circumstances, and it would do the same for cases of dissent and for cases where Jo is undecided. Such a manual contains all the data relevant to the meanings of Jo's sentences. For Quine reasons that it would be enormously problematic to ask for more. To demand more would be to say that even if you knew all of Jo's linguistic dispositions, you would nevertheless not understand her language; there is something you might not know about what Jo means by her sentences. But what you would know about Jo is the most anyone knows about anyone, so far as language is concerned; we learn language by observation of others, observing what people say when. Such is Quine's behaviourism, the aforementioned requirement that linguistic phenomena be at least potentially public.

It should be possible, then, to construct an account of an individual's language strictly based on such a register. In order to sharpen the scenario, we can imagine venturing into an hypothetical region of the globe where only a previously untranslated language was spoken. We conceive the task, as Quine does, as that of devising

a systematic translation into English. If we assume, as we do, that we understand English, then this is to give a manual which pairs each native sentence with an English sentence which is appropriate in an equivalent circumstance. Such is the task of **radical translation**.

To begin, we might watch the natives going about their business, watching out for speech, and especially for hints as to what they are talking about. Suppose it so happens that a rabbit appears, and a native cries 'Ga-va-gai!' (Quine's famous example). This could be a sentence equivalent to 'Rabbit!' – in the use where it is equivalent to 'There's a rabbit!' – it could be a name of that particular rabbit, or it could be something else entirely, whose connection with the appearance of a rabbit on this occasion is purely contingent; maybe the appearance of a well-fattened rabbit prompted the native to remark on how good dinner is likely to be – or thoroughly accidentally – the native was talking about the recent rainstorm, and the rabbit just happened to appear at that precise moment. So you try it, while the rabbit is still visible: 'Gavagai?', you ask; 'Ja', says the native. A little later, when the rabbit is gone, you try again: 'Gavagai?', you ask; 'Nie', says the native. Here, then, is an hypothesis: 'Gavagai' is translatable, perhaps, as 'There is a rabbit' or something like it, and 'Ja' and 'Nie' are signs of assent and dissent, the equivalents of 'Yes' and 'No'. Suppose you then test this hypothesis, trying it on various speakers, in the presence or absence of rabbits, and that the hypothesis is borne out.

Quine strives to describe the natives only in causal terminology – being caused to utter certain sounds in certain observable circumstances and so on. That is the sort of thing he wants to find out. He does not describe the speakers in terms of assertion, as saying things that have truth-conditions, or indeed as referring to objects.

'Gavagai' is what Quine calls an **observation sentence**. The disposition to assent to an observation sentence comes and goes depending on what is apparent in the immediate observable environment; such is what Quine once called the *intersubjective stimulus-meaning* of an observation sentence. These sentences are not *subjective* as in 'My stomach hurts', which pertain to things not observable publicly. And they are not what Quine calls **standing sentences**: these are normally inter-subjective or public, but once established as true or as false, do not change their truth-values. Examples in English include 'Neil Armstrong walked on the Moon in 1969' or 'A helium atom has two neutrons in its nucleus'. It is obvious that what you are looking for at the beginning of radical translation is observation sentences like 'Gavagai', not subjective sentences and not standing sentences; as Quine says, observation sentences constitute the 'entering wedge' of translation. A native's disposition to assent to 'Gavagai' will come and go depending on the comings and goings of rabbits. But since, for example, a helium atom always has two neutrons in its nucleus, a native sentence translatable as 'helium atoms have two neutrons in their nuclei' cannot be discovered by watching for correlations between dispositions to assent and changes in the observable environment. Like every other standing sentence, correlations of that kind are irrelevant to their meaning. They all have equivalent stimulus-meanings, because they all tend to retain their truth-values whatever is happening in the environment. Stimulus-meaning does not add up to meaning, by any stretch.¹

In order to progress beyond the observation sentences, you have to begin to disassemble the sentences, assigning translations not to sentences but to smaller units or expressions. You know anyway that merely translating sentences one by one would never suffice for the translation of the whole language; the idea is to provide a means for translating *all* the sentences of the native language, of which there are in principle infinitely many. This is where the idea of compositionality, or, rather, what Quine calls a system of 'analytical hypotheses', sets in. A simple example is that if we translate, say, 'Gavagai blei' as 'There is a white rabbit' and 'Bollogai pig' as 'There is a black snake', an obvious guess at translation is that 'blei' goes with 'white' and 'pig' goes with 'black'. If a black rabbit should be available, you can try out 'Gavagai pig' on the natives. Quine terms this procedure 'analogical substitution'.

Those are still observation sentences, but the translation of individual terms is the key towards the translation of standing sentences. Indeed there is no other way, no other sort of evidence for their translation (except general considerations relevant to any theory, such as simplicity). Essential to translating those will be the translation terms for logical relations, the equivalent terms for 'if-then', 'all' and the like. Suppose for example we suspect that 'og' is a sentence-connective of this language. If, for any translated observation sentences # and %, a representative native assents to '# og %' if they assent to both # and % individually, and otherwise dissent from '# og %', then one can be fairly sure that 'and' translates 'og'. Similar strategies are employed for other logical connectives; we can thus extract the logical skeletons underlying the imaginary language and thus cope with such standing sentences and, indeed, theoretical sentences as that whose English translation is 'If it's snowing, then it's cold'.

But we shall not pursue this further, as we are in a position now to understand a key Quinean lesson.

● INDETERMINACY

Even if 'Gavagai' can be translated in this way as 'Rabbit!', it by no means follows that there are not other translations which, intuitively, are semantically inequivalent – that is, which do not mean the same. The observation sentences 'Rabbit' and 'Rabbit-stage!' – the latter reports a momentary temporal stage of a rabbit – or 'There's a rabbit!' and 'There's a rabbit-stage!' are associated with the same observable circumstances; one is assertable just when the other is. Further 'stimulus-synonymous' sentences are 'There's an undetached rabbit-part!', 'Rabbit-hood is manifested!' and other more artificial things. Thus the fact that 'Rabbit!' (or 'There is a rabbit!') and 'Gavagai!' are equivalent *sentences* in this respect does not imply that 'rabbit' and 'gavagai' are synonymous or co-extensive *terms*. In particular we have to find out whether the term 'gavagai' has the same *reference* as 'rabbit' or 'rabbit-stage'; and to find that out, we need to translate some native expression as 'is the same as' or equivalent.

Suppose we have identified the native construction 'ipso' as a candidate for 'is the same as' and 'yo' as a demonstrative pronoun, like 'that' in English. The native, we find, affirms:

(1) Yo gavagai ipso yo gavagai.

He affirms this, we find, when and only when we point at the same rabbit both times. This would seem to bear out the hypothesis that 'gavagai' means rabbit rather than rabbit-stage. But it doesn't. Here are two translations of the sentence (1):

(2) That rabbit is the same as that rabbit.

(3) That rabbit-stage is part of the same animal-history as that rabbit-stage.

(2) and (3) are correlated with exactly the same observable circumstances; we have the same rabbit if and only if we have rabbit-stages that are part of the same animal-history. In English, (3) is more complex than (2), but the two are potential translations of the same sentence of the native language. So the native's speech dispositions do not seem to fix the reference of the term 'gavagai'.

Of course, the expressions involved in (1) – 'gavagai', 'ipso' – have uses in the rest of the native language, so the translations of these will have ramifications for translations of other sentences. But just as the data left us with choices in assigning references to the parts of (1), so these choices can be compensated for where other choices emerge in connection with other parts of the translation manual. This is not to say we or the natives do not know the difference between rabbits and rabbit-stages – as long as we are speaking a language such as English, 'rabbits = rabbit-stages' remains false; but, all the same, we can translate the word of the imaginary language in either way, so long as corresponding adjustments are made elsewhere.

This does not mean that translation is impossible; it means that translation is too *easy*, with too many right answers. Beyond the parameters set by linguistic behaviour, there is nothing to be wrong or right about. In practice, one will simply select the option that is most familiar – one can just stipulate that the term 'gavagai' is equivalent to the term 'rabbit', but a cagey translator knows that doing so is only a convenience, that in so far as the facts are concerned one might as well have chosen another from among the alternatives. And, who knows, maybe to the native way of thinking, another choice *would* have seemed more natural.

Such is the simplest argument for what Quine calls the *inscrutability of reference*, part of his broader thesis called the *indeterminacy of translation*. There are several further arguments and implications we might consider – indeed stronger arguments and more radical implications we shall consider a bit later – but this is enough to bring out the central point: the notion of proposition, of sameness of meaning, has no determinate application to the plain facts of the use of language, but is *not needed for an account of such facts*. Pre-theoretically, perhaps, we think one expression correctly translates another just in case it has the same *meaning* as the other; a sentence correctly translates another just in case it expresses the same proposition as the other. We think that two translation manuals, insofar as they are successful,

converge: we expect them to deliver the same translations (allowing, of course, for discrepancies on points of emphasis, style and other grammatical and lexical alternatives that, as we say, amount to the same thing). Quine denies this. Even if you and I were to go about our respective tasks with unimpeachable correctness, there is no reason to assume that we must devise manuals that converge in this way. The two manuals could correctly translate one sentence of the native language into different sentences of English that 'stand in no plausible relation of equivalence, however loose', as Quine puts it (1960: 27). The assignment of particular meanings to expressions is irreducibly intuitive and interest-relative, not something that could be validated by the impersonal procedures of science. Indeed, if we find that two inequivalent verdicts are respectively each part of a complete analysis of a given person's language – an analysis that painstakingly catalogues all the person's linguistic dispositions – then the only way to maintain that the two verdicts herald different ascriptions of meaning is to suppose that the differences of meaning are real but do not show up in the person's linguistic dispositions. But to deny that, as noted earlier, is basic to Quine's naturalism; he holds not only that it assumes an unfeasibly high standard, but that it's unclear what could possibly meet it.

Since language is in Quine's estimation public, this has immediate ramifications for one's own grasp of language. It's natural to think that Quine's conclusion is that although there is no saying what the sentences of the foreign population mean, one's grasp of sentences in one's own language is unaffected. However, since language is a public phenomenon, not a private phenomenon, one's grasp of meaning is necessarily something corroborated in intercourse with others. This shows up when we reflect that if the two unlike translations-into-English are T and T*, then these would amount to different ways of putting the same information in English. Perhaps T is more intuitive or natural in English, but if you spoke in the T* way, it might seem odd, but that is all. 'Radical translation,' says Quine, 'begins at home'; you cannot claim *really* to be referring to rabbits rather than rabbit-stages, even though, overwhelmingly, the word 'rabbit' is common and 'rabbit-stage' is comparatively rare, and as a surface point of English usage, one says that one is referring to rabbits.

● MEANING AND ANALYTIC TRUTH

According to Quine, there are no facts that would justify a science of meaning, not one which assigned particular propositions to sentences and objects to singular terms. However, Quine is well aware that the word 'meaning' is a well-established word of ordinary language; we do ordinarily speak of the meanings of words, and, indeed, we have such things as phrase books and dictionaries, such people as lexicographers and translators. Quine does not think all this is a sham. He thinks that at least much of what we say when we speak of the meanings of words can be explained in terms of linguistic dispositions or linguistic behaviour: one mentions a word or phrase that one would be willing, in the cited circumstance, to use in place of the word or phrase in question. One could just as well say 'In place of that, I would be disposed to say ...' rather than 'It means ...' As part of this general

attitude, he explains how in fact the word 'meaning' is employed by ordinary people as well as in a more informed and systematic manner by professional lexicographers, and he finds a miscellany of different uses that do not add up to anything like a systematic theory. In order to 'give a meaning', sometimes one cites an equivalent word or phrase: one that is equivalent for certain purposes or another that's equivalent for different purposes; at other times, one directs one's listener or reader not to another form of words but to the thing itself, as in "'Tiger" is a word used for those large Asiatic cats that normally are orange with black stripes'; at other times, we get the point across by means of analogies, hints or a list of examples of correct usage. It is a profoundly useful but ultimately imprecise service, not so much a science as an art.

Since at least the days of Hume, philosophers have spoken of an important division of statements into those that express 'matters of fact' and those that express 'relations of ideas'. Kant introduced the terminology that persists today, distinguishing 'synthetic' statements from 'analytic' ones. A modern definition, one that explicitly takes the things said to be analytic or synthetic to be *statements* – or, more strictly, sentences-at-contexts-of-utterance – rather than propositions, judgements or thoughts-in-the-mind, would be: a statement is *analytic* if and only if it is a logical truth or can be transformed into a logical truth by exchanging synonyms (see the Introduction). Consider, then, 'No bachelor is married'. As it stands, this is not a truth of logic; it is of the logical form 'No A is B', which has plenty of instances, such as 'No human is on Venus', which are true but not logically true, and also such instances as 'No human is in California', which are not logically true because they are false. But 'bachelor' is synonymous with 'unmarried man'. So 'No bachelor is married', by substituting synonyms, can be transformed into 'No unmarried man is married'. And that is a logical truth, an instance of the logical form 'Nothing not-B and C, is B'.

We can characterise synonymy as two expressions having the same meaning. As we've seen, Quine, although he agrees that the notion of meaning is useful for ordinary purposes, denies that the notion of meaning has a more exacting use, that for precise scientific and philosophical purposes the notion is best left aside in favour of the notion of linguistic dispositions. Quine is thus very much against the idea that analyticity should play a serious and indispensable role in philosophy or science.

● THE ARGUMENT OF 'TWO DOGMAS OF EMPIRICISM'

In fact, Quine attacked the notion of analyticity *before* directly attacking the notion of meaning. The original attack on analyticity was only indirectly an attack on meaning (his famous or infamous article 'Two Dogmas of Empiricism' appeared in 1951 – it appears slightly revised in Quine 1961 – and the book *Word and Object* appeared in 1960). It was Carnap's view that was Quine's primary and explicit target, but also in the picture were Ayer's view, the views of Strawson and Grice, and of many others who depended on the distinction that Quine sought to undermine (in 1956, Grice and Strawson co-wrote a famous reply to Quine, 'In Defense of a Dogma').

In Quine's article, he argues that no definition or characterisation of analyticity is possible that does not presuppose the legitimacy of the notion of meaning or notions very close if not identical to meaning such as *conceptual content*, *verification conditions*, *propositional content* or Fregean *sense*. But this might leave one puzzled if one thinks that some such concept is all right, even if rough around the edges. The key is that the overarching question for Quine is one about the nature, structure and dynamics of knowledge. We can sketch a caricature of the view that Quine attacked, one that is not, however, misleading on the fundamental points.

There are statements about what is immediately given in sensation or experience; then there are statements that are not themselves experiential but which are conceptually related to those that are. The nature of the linkage is analytical. It is because of various complex analytic relations to experiential statements that a non-analytic statement *S* – an ordinary one like 'There is beer in the refrigerator' or 'She's single', or a theoretical statement like 'Squirrels do not live as long as 25 years' or 'Ethanol freezes at -114 degrees Celsius' – comes to have the particular meaning that it has, and will logically entail certain experiential statements *E*₁, *E*₂ and so on. Those statements count collectively as an experiential test for the veracity of statement *S*; if *E*₁, *E*₂ ... are all true, then *S* must be accepted as true; if at least one of *E*₁, *E*₂ ... is false, then *S* must be rejected as false. Inevitably, then, some statements will be confirmed *irrespective* of experience; an obvious case will be 'Either there is beer in the refrigerator or there is no beer in the refrigerator'. These are the analytic truths, including 'No unmarried men are married' and 'No bachelor is married'. Such is known as *epistemological reductionism*: the justification of statements always proceeds from a foundation in experience upwards to theoretical statements that cannot in practice be compared directly with experience, but each of which is nevertheless equivalent to some perhaps complicated combination of statements which can be. Every cognitively meaningful statement is either analytic or is equivalent to a combination of verifiable statements.

Quine's idea is to replace this picture whereby justification is *linear* with what is known as epistemic **holism**. Suppose a statement *S* is up for testing, and a correspondingly expected experiential statement *E*₅, one that is entailed by the theory as it stands, is *falsified*. Typically, statement *S* will entail the truth of *E*₅ not by itself but only in concert with many other statements: statements pertaining to the nature and reliability of the testing equipment, the chemical composition and behaviour of various substances, logic and arithmetic and so on. In principle, then, what we test is a whole raft of statements, not just *S* in particular. It's just that normally the other members of the raft will be more firmly established than *S* itself, for in an actual case we are testing *S* because it is more critical to an ongoing theory than the others. But nevertheless, in principle one could, in the case where *E*₅ proves false, withdraw not *S* but some other member of the raft of statements which together with *S* entailed *E*₅. And those statements will, in turn, be connected with others, and the others with others and so on. It's very much an 'in-principle' point, but in principle we are testing the whole of science with each particular test, experiment or observation – including the so-called 'analytic statements' and the statements of mathematics. Yes, even the statements of arithmetic or mathematics, generally, are

revisable: in certain, dimly glimpsed and exceedingly improbable circumstances, we could adjust them.

In a famous passage, Quine writes:

The totality of our so-called 'knowledge' or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges. Or, to change the figure, total science is like a field of force whose boundary conditions are experience. A conflict with experience at the periphery occasions readjustments in the interior of the field. Truth values have to be redistributed over some of our statements ... [T]he total field is so underdetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to reevaluate in the light of any single contrary experience. No particular experiences are linked with any particular statements in the interior of the field, except indirectly through considerations of equilibrium affecting the field as a whole.

(Quine 1961: 42–43)

According to holism, the sum of human knowledge, speaking a bit figuratively, is a loose-jointed web of belief; it is not a rigid, vertical structure built upon a foundation of experiential statements, with analytic statements providing the cement. Knowledge can be characterised without making use of the concept of analyticity or the concept of meaning. The problem Quine was attempting to solve is an epistemological problem, not a problem about linguistics, but Quine's solution to the epistemological problem had a direct effect on linguistics; and, indeed, the period of writing 'Two Dogmas' saw Quine branching out from logic and the philosophy of mathematics to the philosophy of language.

● REGIMENTATION; ANALYSIS VERSUS REPLACEMENT

We shift now to the task of explicitly specifying the most general shape, fabric and contours of the language which Quine envisages as adequate for expressing, in the most rigorous way possible, the sum of human knowledge. On this whistlestop tour, we can only glimpse the topic fleetingly, but it is of some value, nonetheless. The connection to his general account of language is simple: the task is to describe the language of a maximally rigorous statement of our theoretical knowledge in terms which are available to typical users of language.

As we have seen, Quine rejects the idea of the meaning of an individual sentence and thus does not believe there is a well-defined task called conceptual analysis. His aim is never to discover, isolate and investigate the concepts or meanings expressed by terms of ordinary language; strictly speaking, he denies that there are any such things. Instead, Quine's remarks on these sorts of issue are driven by his aim to

fashion *replacements* of certain imprecise or otherwise problematic ordinary locutions with theoretical language that is free of those defects – the aim not being linguistic reform for its own sake but to provide a rigorous mode of expression for the statement of one's official theory of a given subject-matter. Some expressions are ordinarily useful, but when pressed harder they fall apart, and often if used they encourage one to make distinctions where in reality there aren't any. Such distinctions can be maintained only in a manner that is arbitrary or misleading, resting upon ideas that are thoroughly unscientific and may lead to contradiction in scientific contexts. For example, the notion of *tall* can be replaced with *taller than*, thus obviating the spectre of a Sorites-paradox arising in our official statements. The notion of a *hill* is useful outside science but is too vague for inclusion within it (for example, the question 'How many hills are there in Montana?' does not have an exact answer). The notion of the meaning of a sentence can be replaced for theoretical purposes by that of a linguistic disposition. (We'll see more of this in Chapter 14.)

A more provocative example is the entire category of proper names. They are rife in ordinary speech, but they are notoriously difficult to account for without considerable philosophical cost, as we have seen (in Chapter 3). Quine's solution is that they can be replaced – not analysed – in a language free of them. The important aspect of their ordinary use can be duplicated by a language without them. If, for example, we have a particular interest in having a 'name' that would serve as the ordinary name 'Socrates' does, we can simply invent a predicate – '*a* Socratises' – stipulating that it is to be true of just that particular man. Then whenever we want to use the name 'Socrates', we form the definite description 'the Socratiser', defining it in Russell's way; clearly, 'Socrates is Greek' will be true so long as 'There is an *x* such that *x* uniquely Socratises and *x* is Greek' is true. Crucially, if for whatever reason the description in fact fails to denote anything, such statements will simply be false, not lacking truth-value. It is plainly artificial, but remember that the aim is replacement, not analysis (and here there is no commitment to Russell's principle of acquaintance). Thus the only singular terms needed are the quantificational variables used to express generalisations, such as 'There exists an *x* such that *x* is Greek', or 'For all *x* and *y*, if *x* is adjacent to *y* then *y* is adjacent to *x*'.

Another point is that Quine does without the ideas of *properties* and *relations* as the meanings of predicates. This is a view which goes back to Quine's early 'On What There Is' (1948), and is given more force in the essay 'Necessary Truth' of 1963. There is on the one hand difficulty in coming up with exact 'criteria of identity' for what he called 'attributes' (Russell called them 'universals'); and on the other hand, it is not necessary to appeal to attributes (properties and relations) to describe the semantics of predicates. For example, '*x* is red' is true of an object just in case the object is red. A given language contains only finitely many simple predicates, so there are no barriers to such a strategy.

This is, I said, merely an impressionistic presentation, but it is enough to see what Quine is driving at. The language of logic – for Quine, the first-order predicate calculus – *without* names except for variables, with the predicates selected in such a way as to be rid of the sorts of defects discussed two paragraphs back, and explained

positively in the way of the last paragraph, is sufficient for the expression of scientific knowledge. Not that one has to formalise one's theory; so long as one's terms admit of ready formalisation; that is not necessary. This is Quine's idea of **regimentation**: the rigorous and streamlined presentation of the whole of one's science, couched in logic-ready terms as just described.

Quine's case that the logic need only be first-order, by the way, is complicated, but two points are worth stating. First, Quine accepts that mathematics is endemic to science as a whole, and accepts set theory as a basis for mathematics; yet set theory can do much of the work of higher-order logic, at any rate all that genuinely need be done, in Quine's estimation. Second, Quine came to be impressed with the fact that unlike second-order logic, first-order logic is 'complete', that is, that within a properly constructed proof-system, all semantically valid schemas are provable. Quine does not see the need for a more inclusive definition of 'logic'.

● FURTHER DISCUSSION: ONTOLOGICAL RELATIVITY, AND INSCRUTABILITY AND INDETERMINACY DISTINGUISHED

Suppose we have succeeded in regimenting the whole of science (of course we have not said precisely what is intended by 'science', but think first of paradigm cases such as physics, biology etc.). It will have terms for physical objects, terms for numbers, terms describing spatial relationships, terms for mass and velocity, terms for chemicals, for protons and electrons, and so on. Now we define the notion of a 'proxy-function'. If for a two-place predicate 'R', there is always a unique object b for any a such that Rab , then we can speak of the corresponding function f : $f(a) = b$. The object a in this case is called the 'argument', and b is the corresponding 'value' of the function. If, in addition, f is 'one-to-one' or 'bijective' (no two arguments have the same value), and for an appreciable portion of the domain of the function (the set of arguments of the function, which is simply the entire universe), the function delivers as value a different object, then f is a *proxy*-function. For a simple example, 'the unit set of' is a proxy-function, i.e., $f(x) = \{x\}$. Another example of a proxy-function would be the cosmic complement of x : the cosmic complement of the San Francisco Bay Bridge would be the entire physical universe with a certain bridge-shaped hole in it (for abstract objects as the arguments, $f(x)$ can be x itself). We'll use 'cosmic complement of' as our proxy-function. Now consider any true sentence which ascribes a single predicate to an object (for ease of exposition we'll pretend we have genuine singular terms). Suppose it's 'The GG Bridge is red'. Now consider the cosmic complement of the GG Bridge. *This* is the cosmic complement of a red thing (the bridge). In general, if Fa is true, then the proxy of a is the proxy of an F . A slightly more complicated version of the same goes for two-place predicates, three-place predicates and so on.

So what? If a theory is taken to have a certain universe, we can, by deploying this trick of proxy-functions, exchange the universe for another, and the truth-values of

all sentences will remain the same. Since there is no other way to get at the references of terms except via the truth-values of sentences in which they occur, it seems to follow that reference is radically inscrutable, in a more sweeping way than what was suggested by the rabbits.

From the point of view of a general reckoning of scientific knowledge, it further appears that *ontology* – what exists – is inscrutable; or rather, as Quine put it in the eponymous essay of 1968, the situation obtains of *ontological relativity*. It is relative to a scheme of reference: by 'the GG Bridge' one habitually says that one refers to the GG Bridge, but we can with the help of a proxy-transformation reinterpret the whole of the language so that the referent of 'the GG Bridge' is the proxy of the bridge. It might be awkward but one would be making no scientific blunder. Indeed, one could adjust one's usage to erase to awkwardness that would arise in communication with normal speakers – saying 'the GG Bridge' when one thinks 'the proxy of the GG Bridge', and no one need know; one might even forget about it.

Quine's conclusion is that reference and ontology are not foundational to human knowledge; it is the structure of predicates in which its factual content really lies.

This is related, in virtue of their both operating at a very abstract level, to another argument for the indeterminacy of translation, which earlier was mentioned but not spelled out. For this we need the notion of the empirical content of a theory, which in turn requires the notion of the observation categoricals of a theory. It is familiar from the basic philosophy of science that an elementary method of observational inquiry is that one contrives or waits for a condition or circumstance, and makes an observation of some kind. Translated in the terminology of Quine and generalised, our interest is in sentences held true of the form 'Whenever O_1, O_2 ', where O_1 and O_2 are observation sentences. These are observation categoricals; a basic example is 'Whenever smoke, fire'. The sum total of these that are held true, then, constitutes the empirical content of a theory. It is one basic job of the rest of theory to imply them.

It is common amongst epistemologists and philosophers of science to hold that non-trivial theories are underdetermined by their evidence. That is, in Quine-speak, they are underdetermined – not determined – by the observation categoricals they imply. Call this the thesis of UD. Suppose E is the empirical content of our theory T_1 , where T_1 is a set of theoretical sentences. UD means that in principle, E could be generated exactly by some other theory T_2 . Now suppose that E is the empirical content of some alien theory, T_A as we have translated it; then by UD, there is some other alien theory T_B which accounts for E equally well. But there is no access to theoretical sentences except via their empirical consequences, that is, via the observation categoricals, indeed by translating the observation sentences they involve. So the question of whether the aliens hold T_A or T_B makes no empirical sense. Observing the natives, by hypothesis, can never choose between ascribing T_A or T_B to them. Whereas the question of what theory we ourselves hold is answered trivially by reciting theory or by pointing to books, the question of which theory the aliens hold can be settled only by translation, but that is just to answer the question of which theory we are thinking of them as holding. The matter is therefore indeterminate; there is no fact of the matter.

In an article of 1970, Quine distinguished sharply the inscrutability of reference from the indeterminacy of translation. The argument at the beginning of this section for the inscrutability of reference involves different reference schemes while preserving the truth-value of every sentence. The indeterminacy of translation, by contrast, involves more: a theoretical sentence might be translated as *p* according to one translation of the whole language but as not-*p* according to another. This is unlikely to happen in the case of obvious truths or falsehoods, but there is no obvious reason to rule it out if the sentence in question were one about which the natives, or ourselves, had no strong opinion. Quine gave the argument of the last paragraph in the same article. In later years, Quine came to question the argument, and since he now was distinguishing sharply arguments about reference of terms from ones about the truth-value of whole sentences, he retreated to characterising the indeterminacy of translation as a 'conjecture'. But this does not defeat Quine's project, as will be underlined in a moment.

● THE PLACE OF NATURALISM

The conclusion of indeterminacy does not mean translation is impossible, that the alien mind is unknowable. It's that translation is too easy to sustain the idea that there is any one right way to portray the alien mind. This goes for the minds of our fellows, and indeed for our mind. The idea is that certain mentalistic notions – the mind's grasping propositions, of the mind's having a kind of non-causal contact with things outside it and so on – are from Quine's scientific point of view ill equipped to serve as materials for genuine explanations. I say 'scientific' point of view, but for Quine that covers any domain of knowledge, and any domain of knowledge must submit to the strictures of naturalism. Of course, ordinary knowledge is not without epistemic value or not in any sense knowledge: it just doesn't measure up to the most stringent demands and fails to be rigorous and maximally objective.

There is hope for one who likes naturalism but is loath to accept Quine's negative conclusions for the idea of meaning. Suppose, as many have argued, that translation in Quine's sense is not indeterminate as he says (and the late Quine does not deny this possibility, as related in the last section). Staying within the naturalistic constraints on Quinean radical translation, suppose Quine has missed certain facts that determine translation. In that case, a notion of the meaning of a sentence would readily be available that would not violate Quine's naturalism: take all sentences in which a given sentence can be correctly translated – i.e. are synonymous. These make up a set of sentences of intuitively equivalent translations (an equivalence class under the relation *x* correctly translates *y*). This set can perform the function of what previously we were calling the meaning, proposition or sense expressed by the sentence: as a vehicle of truth, as what, by definition, all sentences that are synonymous with the original sentence have in common, and as the object of propositional attitudes. Call the set a meaning*. Thus Quine's picture could conceivably be brought into a closer harmony with common sense as well as with the idea of a theory of meaning. Nevertheless, the explanation of *why* the classes are as they are would be independent of, and prior to, the application of the concept of meaning*. The concept of meaning* would still be without causal robustness and would be incapable

of serving in genuine explanations: fundamentally, it's not *because* a sentence has a certain meaning* that it plays its part in linguistic behaviour, for saying that it has that meaning* does not add anything to saying that it plays the behavioural role. Indeed, within naturalism, it is completely unclear what role a causally robust notion of meaning could play. To think that it can play such a role would be to commit the sin of private language as discussed by Wittgenstein, as related in Chapter 5.

● HISTORICAL NOTES

W. V. Quine came upon the scene in the first instance as a mathematical logician; most of his published papers and books until the 1950s were on logic and set theory; but meeting Rudolf Carnap in Prague in 1933, who was working on epistemology and the philosophy of language as well as logic, made a lasting impression. He was impressed but had doubts, some of which were expressed in 'Truth by Convention' of 1935, when Quine was all of 26. He established his credentials in set theory and logic with his 'New Foundations for Mathematical Logic' in 1937. After the war – during which Quine put his logical skills to work on finding algorithms for designing more efficient electrical switching circuits – his criticisms of uncritical use among philosophers of analyticity were famously expressed in his 'Two Dogmas of Empiricism' of 1951, together with his sketch of the dynamics of knowledge, which dispensed with the notion. Then came *Word and Object* of 1960, which among other things set out the case for the indeterminacy of translation and provided more in the way of detail for an explanation of human knowledge and language without relying on the concept of meaning. He went on developing and fine-tuning the view through the 1990s, especially with the *Roots of Reference* (1974) and *Pursuit of Truth* (revised edition 1992). More generally, Quine has influenced the course of analytical philosophy deeply in various ways but most conspicuously by setting some of its leading problems and for articulating solutions to them for philosophers to celebrate or criticise. Interestingly, in *Roots of Reference*, he proposed a naturalistically acceptable definition of analyticity, thus allowing the use of the word while maintaining that it could not play the sort of epistemological role envisaged by Carnap and others. There aren't many out-and-out followers of Quine, but recent figures who have been broadly supportive include Gilbert Harman, Daniel Dennett and Peter Hylton. Not exactly a follower of Quine but a PhD student of his was David Lewis (1941–2001), whose work still dominates metaphysics. In Chapter 10 we will see that Donald Davidson, partly inspired by Quine, but who departs from him on the central question of whether a rigorous theory of meaning is possible, developed a view which has gained a more visible following.

● CHAPTER SUMMARY

Quine's criticisms of the work going on in the philosophy of language were motivated by his scientific naturalism, by the position that a serious understanding of linguistic phenomena should not uncritically help itself to the idea of meaning and

related concepts. Closely connected is that linguistic significance must be accessible from a third-person point of view; in particular it must be reconstructable in terms of linguistic behaviour.

The thought-experiment of radical translation involves translation into English of a previously unknown (to the translators) language; this clarifies what naturalism requires and prohibits, and it affords a way of finding out the objective standing of the idea of cognitive meaning, of propositional content. The meaning-free concepts with which translation begins include the concept of an *observation sentence*, which goes from true to false depending on what is apparent in the immediate observable environment, and the concept of a *standing sentence*, which does not change its truth-value once established as true or as false. Quine's conclusion is that radical translation is indeterminate; for example, there is no objective difference between a translation of a native observation sentence such as 'There is a rabbit' and 'There is a rabbit-stage' – for one can make adjustments to the translations of individual expressions to even out any bumps. As mentioned above, the conclusion of indeterminacy does not mean translation is impossible, that the alien mind is unknowable; it's that translation is too easy to sustain the idea that there is any one right way to portray the alien mind. Later Quine separated the indeterminacy of translation from inscrutability of reference; arguments for the latter are stronger.

Analytic truth is truth due solely to the meanings of words. In 'Two Dogmas of Empiricism', Quine dismisses meaning from the arsenal of naturalistic concepts and thus dismisses analyticity. Analyticity played a central role in theories of knowledge from Hume and Kant to Ayer and Carnap. In holism, Quine sketches an alternative theory of knowledge that dispenses with the concept: rather than a false prediction enjoining a specific alteration of theory according to analytic rules, the effect of a false prediction is a more flexible response, with various alterations possible. Human knowledge is like a loose-jointed web of belief, not like a vertical structure built upon a foundation of experiential statements.

It is easy to read some of Quine's work as attempting to *analyse* ordinary expressions, but in fact he proposes *replacement* of parts of ordinary or even scientific language. Ordinary expressions are often vague, imprecise and even contradictory in their implications; Quine's solution is not to find out better what they mean but, if the need is felt for some theoretical purpose, to propose replacements which serve the necessary functions without those problems.

● STUDY QUESTIONS

- 1 If the notion of meaning is *indispensable*, from a practical point of view, how can science take the view that it is ultimately not real?
- 2 Suppose Mary speaks only English, and Bob speaks only a variant of English, called Quinglish. Strangely enough, Bob's language and Mary's have never been in contact, not even indirectly. But then Bob and Mary meet, and set about at once to undertake radical translation of each other, as one does. Bob translates an English sentence S, 'A rabbit is present', as the Quinglish sentence S*, 'A rabbit-stage

is present'. Mary, sensible girl that she is, translates the Quinglish sentence S^* as the English sentence S^{**} 'A rabbit-stage is present'. Therefore, since S^{**} translates S^* , and S^* translates S , S^{**} translates S . But then the English sentence 'A rabbit is present' is equivalent to the English sentence 'A rabbit-stage is present'! How should Mary respond to this?

- 3 Does Quine's theory imply that we don't know our own minds, our own thoughts?
- 4 An alternative way of defining analyticity is this: a given statement is analytic if and only if *accepting* it is a criterion for *understanding* it (that is, in order to understand the concepts involved, one must accept a given statement). Would such a definition fare better against Quine's criticisms?

● PRIMARY READING

The first Quine is inevitably the hardest; I recommend 'Speaking of Objects', in his *Ontological Relativity and Other Essays* (1969), and 'Two Dogmas of Empiricism', in his collection *From a Logical Point of View* (second edition 1961), then the essay 'Things and Their Place in Theories', in his collection *Theories and Things* (1981). Not for the faint of heart is *Word and Object* (1960).

● SECONDARY READING

Hylton, P. and Kemp G., 'Willard van Orman Quine', in the *Stanford Encyclopedia of Philosophy*.
 Kemp, G. (2023) *Quine's Philosophy: An Introduction*.

● NOTE

- 1 In fact, we speak of things happening in the environment only as shorthand. Strictly, the stimulus-meaning of an observation sentence will involve only the actual nerves being stimulated in characteristic ways. For a native sentence whose role is that of 'It's hot!' it will be relatively simple, but for 'Gavagai' it will be extraordinarily complex, taking in all the various possible sightings, glimpses and other sensory manifestations of rabbithood, as well as events at the sensory surfaces that are typically not caused by rabbits but by events which fool the subject into assenting to rabbit when there is no rabbit. But this introduces further complexity, which for our purposes we can avoid.

7

• kripke¹ on naming and necessity

● NECESSITY, POSSIBILITY AND POSSIBLE WORLDS: A PRIMER

It seems to be a *necessary truth* that $2 + 2 = 4$: it could not have been false. Other seemingly necessary truths are that everything is identical with itself, that if x is bigger than y and y is bigger than z , then x is bigger than z .

We know that ice cream exists. But it is *possible*, in the sense that it might have been the case, that ice cream didn't exist. Ice cream might not have existed. It is a truth, but not a necessary truth, that ice cream exists.

What are necessity and possibility? For present purposes, we are going to assume what is most convenient, namely the interpretation in terms of **possible worlds** (qualms about them will be addressed shortly). A necessarily true proposition is a proposition that would have been true *no matter what*. It *could not* have been false. However the world had gone, the proposition would have been true. Thus: the proposition is true in *every possible world*. Similarly, a possibly true proposition is one that *might have been the case* or which *is the case*. The proposition is true in *some possible world*. So we have:

A proposition is **necessary** if it is true in every possible world.

A proposition is **possible** if it is true in some possible world.

If a proposition is necessary then, trivially, it is possible. The actual world is one of the possible worlds: if a proposition is true in the actual world then, trivially, it is true in a possible world. The actual world is the way things actually are.

If p is a proposition that is possible but not necessary, then p is **contingent**; if it is actually true, like the proposition that ice cream exists, then we can speak of it as a *contingent truth*.

The words ‘possible’ and ‘necessary’ are not always used in this way, especially not outside of philosophy. So let us be clear. We are talking about what in philosophy is standardly called *metaphysical modality*. It is a **metaphysical** notion, not an **epistemological** notion. In speaking of a proposition being possible we mean either it *is* true, or it *could have been* true. We don’t mean, ‘I don’t know that it’s not true’. For example, we *know* with certainty that the highest mountain on Earth is not in Tahiti, but it *could* have been. In some possible world, Earth’s highest mountain is in Tahiti. It *might have been the case* that the highest mountain on Earth is in Tahiti, even though it isn’t and we know that it isn’t.

Similarly, by ‘necessary’ we don’t mean ‘it is certain’. Outside philosophy, we do say ‘not necessarily’, meaning ‘we can’t be certain that it is so’, but that is not the sense we have in mind. This is clear from the fact that some necessary truths are not certain, not even known. For example, truths of mathematics are standardly taken to be necessary truths, but not all of them are known.

The traditional epistemological analogues of necessity and contingency are the *a priori* – knowable by reason independently of experience – and the *a posteriori* or *empirical*, knowable only through experience.

For convenience, we sometimes write:

‘Nec (p)’ for ‘it is necessary that p’.

‘Pos (p)’ for ‘it is possible that p’.

‘Nec’ or ‘it is necessary that’ is a *sentential operator* (a sentential connective), like ‘it is not the case that’.² But unlike ‘it is not the case that’, ‘Nec’ is not *truth-functional*; it is *non-extensional*, in the sense introduced in Chapter 2 – and in particular an **intensional operator**. The truth-value of

Nec (p)

is not determined by the truth-value of p. For example, the truth-value of ‘Nec (the Earth is round)’ is not settled by the fact that the Earth is round. ‘Nec’ thus differs from truth-functional sentential operators such as those appearing in

it is not the case that p

p and q

p or q

In these cases, the truth-value of the whole is determined by the truth-values of the constituents – by that of p in the first example, and by those of p and q in the others.

Possibility and necessity are generally held to be *inter-definable*, in the sense that each can be defined in terms of the other:

A proposition is necessary iff its negation is not possible.

A proposition is possible iff its negation is not necessary.

What is a (non-actual) possible world? Is it like the actual world except for being spatio-temporally discontinuous from it, as David Lewis famously held? Or is it unlike the actual world in being abstract rather than material or concrete? Or are we to take the phrase ‘possible worlds’ merely as an heuristic, that it is just a ‘manner of speaking’ and possible worlds don’t *really* exist at all? It is best not to worry too much about this (you can worry about it in your metaphysics class). A possible world is simply a way things might have been: a way things could have been. If you don’t like the air of science-fiction attending the phrase ‘possible world’, you may call them possible *situations* or *counterfactual circumstances* (except you cannot call the actual world a counterfactual circumstance, since the actual facts are not contrary-to-fact). Or if worse comes to worst, you can mentally take ‘Nec (p)’ as explaining ‘p is true in every possible world’, etc.

We can take any proposition and possible world – any way things could have been – and ask whether the proposition is true *at* that world. More generally, we can ask for the truth-value of a proposition at all those worlds at which certain specified conditions hold. This just means: consider these possible circumstances; would that proposition have been true under those circumstances?

Possible worlds make it straightforward to interpret not only ‘Nec’ and ‘Pos’ but also **counterfactual conditionals** (*subjunctive* conditionals). Consider:

If Jones had been wearing a seat belt, he would have survived the collision.

We can say this appropriately only if the antecedent is false – Jones was *not* wearing a seat belt. Thus, if we interpreted this as a material, truth-functional conditional, it would automatically come out as true. As would:

If Jones had been wearing a seat belt, he would **not** have survived the collision.

For the truth-table for the material conditional ‘ \rightarrow ’ has it that such conditionals are true whenever the antecedent is false. This is the same as: ‘Either Jones was not wearing a seat belt, or Jones survived the collision’, which is true, since Jones was not wearing a seat belt. So how do we determine whether it is true that if Jones *had been* wearing a seat belt, he *would have* survived the collision?

Easy: consider a possible world that is just like the actual world except that Jones, contrary to fact, was wearing a seat belt at the moment of the collision. Keep everything else, or as much as possible, the same. This is, as David Lewis puts it, the *nearest* possible world to the actual world at which the antecedent of the conditional holds. It’s the *closest* one. Does Jones survive in that world? If yes, the counterfactual conditional is true; if no, it is false.

● THE DESCRIPTIVIST PARADIGM

Both Russell and Frege can plausibly be regarded as having held a *descriptivist theory of proper names*. For ordinary proper names, such as ‘Aristotle’ or ‘Madonna’,

both figures hold that on each occasion of using or understanding the name, what the speaker or understander means or understands by the name can be expressed by a definite description.³ In this sense, (ordinary) (particular uses of) proper names are strictly equivalent to definite descriptions: they *abbreviate* them: they function as *shorthand* for them. Thus suppose that Beethoven reads the following sentence:

Mozart is dead.

What he grasps, when he reads and comprehends this sentence on this occasion, will be something like: *The composer of Don Giovanni is dead*. If so, then at that moment, the following two sentences will for Beethoven be synonymous; they will express the very same proposition, at least at that moment:

Mozart is dead.

The composer of *Don Giovanni* is dead.

True, Russell and Frege differ on how a definite description such as ‘the composer of *Don Giovanni*’ is to be explained; Frege holds it to be a genuine singular term whose job is to *refer*, whereas according to Russell’s contextual definition of definite descriptions, it contains a quantifier – ‘the’ means ‘exactly one’ – and therefore does not strictly speaking refer. However, this difference will not matter for what follows. The view seems to answer the question: in virtue of what facts about one’s understanding does one talk about an object by using a name? But the view has serious drawbacks, especially now with modality in the picture.

● KRIPKE’S OBJECTIONS TO THE DESCRIPTIVIST THEORY OF PROPER NAMES

Suppose you, like Beethoven, at a certain moment understand ‘Mozart’ as *the composer of Don Giovanni*. In that case, you will understand the following sentence as *analytic*, true by virtue of meaning (make sure you’re clear on the topic of analyticity; see the Introduction, p. 6–7):

(1) Mozart = the composer of *Don Giovanni*.

In fact, this sentence, understood in that way, will be strictly synonymous with both of the following:

(2) The composer of *Don Giovanni* = the composer of *Don Giovanni*.

(3) Mozart = Mozart.

Kripke points out that since all analytic truths are necessary truths,⁴ (1) must be a *necessary truth*, according to descriptivism. It would also be *a priori*: since obviously

(3) and hence (2) are *a priori*, and since (1) expresses the same proposition as (2), (1) is *a priori*.⁵

The problem is obvious: (1), Kripke observes, is *not* a necessary truth. It might have been the case that Mozart never wrote that magnificent opera; of course, he did write it, but he *could have* stuck to instrumental music; he could even have been run over and killed by a carriage as a small child, and never written any music at all. There are possible worlds in which (1) is false.

Nor is (1) *a priori*: that (1) is true is a piece of empirical (*a posteriori*) knowledge. Of course, we can know (3) with perfect certainty, but we can intelligibly doubt the truth of (1) – we have very good evidence that it is true, but we don't know it with *perfect* certainty, not in the way that we know the truth of (3). The falsity of (1) is conceivable. So (1) and (3) must have different *cognitive values*, and hence do not express the same proposition. The very sort of argument – from cognitive value – that Frege used to gainsay the naïve theory can be used against his alternative. Nor is (1) *analytic*: it is not because of the *meanings of words* that (1) is true. It is no triviality like (2) or (3), or like 'every square has four sides'.

The descriptivist idea seems wrong from the point of view of *metaphysics*, because it entails wrongly that propositions such as (1) are necessary rather than contingent; seems wrong from the point of view of *epistemology*, because it wrongly entails that propositions such as (1) are *a priori* rather than *a posteriori* (empirical); and seems wrong *semantically*, because it wrongly entails that propositions such as (1) are analytic rather than synthetic.

There is a fourth kind of objection, which some find the most telling. Consider the name 'Josef Haydn'. Suppose you know that this is the name of a composer who lived around the time of Mozart. But suppose you don't know anything more than that; you can't name any of his compositions, and so on. In short, you lack an *individuating* description: a description that would pick out the name's bearer from everything else in the world. You cannot formulate *any* sentence of the form:

Josef Haydn = the F.

– which you know or at least believe to be true. But if so, then according to the description theory of names, *you cannot succeed in referring to Haydn*, not even by using the name 'Josef Haydn'. In that case, if you were to say or think 'Josef Haydn was a composer' then *you do not succeed in saying or thinking something about Joseph Haydn!* But that seems quite wrong: it seems perfectly clear that you do possess some information *about Haydn* – you know him to be a composer who lived around the time of Mozart.

We can call the argument of the last paragraph the *argument from ignorance* against descriptivism. The argument is closely related to the *argument from error*. Suppose George is like most people in accepting just one fact about Dante: he wrote the *Divine Comedy*. According to descriptivism, then, 'Dante = the author of the *Divine Comedy*' is analytic for George. Suppose that in fact it was written by the humble Adriano; he wrote in absolute secrecy until discovered by Dante, who then strangled

Adriano and took credit for the great poem. But no one knows this. The descriptive theory has 'Dante wrote the *Divine Comedy*', as understood by George, wrongly coming out as true, and about the wrong man; surely Dante, even in these unusual circumstances, is the referent of 'Dante', rather than Adriano being the referent of 'Dante'. The fact that propositions such as (1) are contingent and empirical shows up dramatically in this sort of case. The assumption that referring to an object by means of a proper name depends on the ability to pick out the object by means of a uniquely identifying description seems to be false, since we often succeed in using proper names to talk about objects apparently without possessing such a description, indeed at least in some cases in defiance of the fact that the descriptions that one would cite pick out a different object.

The conclusion of these lines of argument may be put by saying that proper names are **directly referring expressions** (following David Kaplan; this is not exactly Kripke's conclusion). That is to say, they do *not* refer by means of some kind of *conceptual representation* of the object such as a Fregean sense (a mode of presentation). Proper names, on this view, have only reference and not sense. In another vocabulary due to John Stuart Mill, they only *denote* their objects; they do not *connote* anything. In another vocabulary due to Ruth Barcan Marcus, they only *tag* their objects. This anti-Fregean line is often called 'direct reference theory' (confusingly, direct reference theory is often called 'Russellian', despite its denial of what Russell said about ordinary proper names).

● RIGID DESIGNATION

A contingently true sentence is true in the actual world, but false in some non-actual possible worlds. For example, 'The Beatles sold more records than the Rolling Stones in 1965' is true in the actual world, but false in some non-actual worlds. A necessary truth, by contrast, is true in all possible worlds.

Just as some sentences vary in truth-value from world to world, whereas others don't, some singular terms refer to different things depending on what world we are talking about, whereas others don't. Consider:

- (4) the number of Martian moons
- (5) the number of positive integers less than 3

Assume that the truths of arithmetic are necessary truths. Both of these expressions refer to the number 2 (the moons of Mars are Phobos and Deimos). However, that is because in the actual world, there happen to be two moons orbiting Mars. Had things been different, Mars would have had more moons or had fewer moons. Thus, with respect to other possible worlds, (4) refers to a different number: in certain counterfactual circumstances, there would have been six moons around Mars, or fourteen, etc. (5), however, refers to the same number *whatever possible situation we are considering*. Since it is a necessary truth that there are exactly two positive integers less than 3, (5) refers to the same object with respect to all possible worlds.

A singular term such as (5), whose referent remains constant from world to world, is what Kripke calls a *rigid designator*. (4) is a *non-rigid* designator (or a ‘flaccid’ designator).⁶

Careful: of course, the *words* ‘the number of positive integers less than 3’ might have been *used* differently. If the history of English had been different, they might have had different meanings. In that case, (5) would have designated a different object. In the same sense, the *words* ‘bachelors are unmarried’ might have meant something else – if ‘bachelors’ meant *porcupines* and ‘unmarried’ meant *hairless* then those words would have expressed a false proposition (*that porcupines are hairless*). But when we ask whether ‘bachelors are unmarried’ expresses a necessary truth, we mean that, *given what the sentence means*, could *the world* have been such that it was false? Could the *proposition* that bachelors are unmarried men have been false? Of course, the answer is no, since ‘bachelors are unmarried’ means ‘unmarried men are unmarried’. Likewise, when we ask whether (5) is a rigid designator, we are asking whether, *given the meaning of the words*, the world could have been such that it designates something other than the number 2. The answer, assuming that arithmetic truths are necessary truths, is no.

Now consider:

- (6) Nixon
- (7) the 37th US president

Richard Milhous Nixon was in fact the 37th president of the United States. Now consider that man, Nixon. Are there worlds in which not he, but someone else, is the 37th president? Yes. So (7) is not rigid. What about (6)? Consider its referent, Nixon. Are there worlds in which not he, but some other man, is Nixon? No! The following is *true*:

- (8) The 37th US president might not have been the 37th US president.

For this says: take the man who (actually) is the 37th US president (Nixon); are there possible worlds in which *that man* is not the 37th US president? Yes.

But compare:

- (9) Nixon might not have been Nixon.

That’s false! We cannot take Nixon, and then find a possible world in which he is not Nixon. There is no world in which *that man* is not Nixon. There is no possible world in which that man is not that man. What is going on is that ‘Nixon’ is a rigid designator: ‘Nixon’ designates the same object with respect to all possible worlds. Sentences of the form of (9) constitute a *test* for rigidity:

For any term put in both blanks, if ‘__ might not have been __’ comes out false, then the term is rigid; if not, the term is not rigid.

Again, the term 'Nixon' could have been *used* differently; Nixon might not have been called 'Nixon'. There are possible worlds in which Nixon is not called 'Nixon'. We are not denying that obvious fact. Nor are we denying that Nixon might have lived a very different life. The idea rather is that when *we* use the word 'Nixon', using it according to its customary meaning, and talk about some possible world or conceivable situation, we are always talking about the same man, namely Nixon, the man actually called 'Nixon'.

The Kripkean points in this section can be summed up as:

- a Genuine proper names are rigid designators (the reverse does not hold, as (5) shows).
- b Genuine proper names refer directly.
- c Directly referring terms are rigid designators, but not all rigid designators refer directly (again consider (5)).

● FIXING THE REFERENCE I: CAUSAL CHAINS

If the descriptivist theory is false, what theory of reference of proper names is true?

Kripke proposes an account that is so simple and plausible to common sense that it's bizarre that it took so long for anyone to think of it. Take our earlier example, 'Haydn'. When you use 'Haydn', you refer to whoever the people referred to from whom you heard the name – friends, authors of books or magazine articles, the internet, people on the radio or television etc. For convenience let us represent those people as a single person, A. Now who did A mean by it? Same thing: A heard the name from some other person, say B. And who did B mean? B got it from someone else, say C. And C got it from D, D from E and so on. There is an **historical chain** of users of the name reaching backwards in time from your use of it back to ... what? Well, eventually we get to people who actually knew Haydn, knew him in person. They were introduced to the man or the boy, and were told 'This is Josef Haydn'. The *first* users of the name were (presumably) his parents: they stood over the gurgling baby and said, 'Let's call him "Josef"'; together with the convention that the child takes the surname of the father; this amounted to a 'dubbing', a ceremony of naming. From that point on, the parents used the name with the intention to refer to that boy; people who acquired the name from the parents then undertook to refer to the boy to whom the parents referred, and thus the chain began. In the economic phraseology that is sometimes used, the parents are the *producers* in the name-using practice, those who inherit the name are the *consumers*.

This theory has the advantage that it explains how it is that even if unbeknownst to you some other person is called 'Josef Haydn', you refer to the composer, not to that other person: who you refer to depends on which chain of communication you connect with.

● FIXING THE REFERENCE II: DESCRIPTIONS

But have we really got away from the use of descriptions? Isn't 'the person they were referring to by N' a description? Yes, but it does not *give the meaning* of the name N. The arguments against the description theory still apply: it is merely a contingent fact about Haydn that he came to be talked about in the way he has been by means of that name.

Thus, assuming still that it was A from whom you got the name, consider the sentence:

(10) Josef Haydn = the person A refers to by means of 'Josef Haydn'.

This sentence is true, so the two singular terms 'Josef Haydn' and 'the person A refers to by means of "Josef Haydn"' in fact have the same *referent*. But they don't have the same *meaning*. They are not *equivalent*: they are not *necessarily equivalent*, not *epistemically equivalent* and not *conceptually equivalent*. They lack the same cognitive content. With respect to a possible world where he is not called 'Josef Haydn', (10) is false: at such a world, it is *not* the case that A refers to Josef Haydn by means of 'Josef Haydn'.

In Kripke's terms, the role of such a description as 'the person A refers to by means of "Josef Haydn"' is not to give the meaning of the name but to **fix the reference** of the name. Frege and Russell held that there must be some sort of intention in the mind of the language-user that determines the referent of a name, and that the *content of that intention* constitutes the meaning of the name. Kripke sees that last move as the crucial mistake: what fixes the reference of a name need not enter into the meaning, the cognitive content, of the name. *The description may be used to fix the reference of the name, but the name does not abbreviate the description.*

We can go further. When successfully using a name, is it necessary to have an explicit intention that fixes the name's reference? No! Once we give up the idea that names must have some kind of expressible cognitive content, all we are interested in is to what the name refers. All we need in the theory of reference, then, is an account of what facts do determine the referent of a name. Put like that, we needn't assume that among these facts must be explicit intentions in the mind of the speaker. We can rather posit a pair of *semantic rules*:

- (i) If A dubs a given object with the name NN, then when A uses the name NN, A refers to that object by means of NN.
- (ii) If C uses a name NN, and C acquired the name NN from B, then C refers by NN to whatever B referred to by NN.

Such semantic rules, such conventions, determine what a speaker refers to by means of a name (even if the speaker is unable to formulate the rule).

Once we recognise how reference-fixing works, we recognise that the referents of proper names *can* be fixed in something very close to what Frege and Russell

envisaged. For example, one might, by the name 'Slowpoke', *stipulate* that one shall refer to the person in the dressing cubicle by using that name. But what one does, in such a case, is merely to specify the object that is to be the referent of the name. One does not thereby specify a *meaning* for the name; one does not accept the name and the description one uses as *synonymous*, as meaning the same. 'Slowpoke', unlike 'the person in the dressing cubicle', is a rigid designator. One can take advantage of the stipulation to say, for example, 'I wish Slowpoke had not gone shopping today!' One makes such a stipulation without doubting that the following is a straightforward contingent truth, not in any sense a necessary truth:

(11) Slowpoke = the person in the dressing cubicle.

So definite descriptions can be used to fix reference: a dubbing or baptism ceremony, of the kind discussed in connection with 'Joseph Haydn', is not necessary. The description, in such a case, might be said to play the role of a pointing-device.

There are examples that seem to cause trouble for Kripke's picture. First, some names plausibly do fit the descriptivist theory. One might think that whatever the facts of pre-historical Greece, 'Homer' as it is used now is actually intended to *mean*, be *synonymous* with, 'the author of the *Iliad* and the *Odyssey*' – it is not merely, as was just suggested, that that description is used to *fix the reference* of the name. Similarly with names like 'Jack the Ripper'; perhaps the name really does abbreviate a certain description – perhaps it is plausible to say that *Jack the Ripper might not have been Jack the Ripper* (whereas it is false that Nixon might not have been Nixon). Second, we've said nothing about proper names for *fictional* entities such as 'Superman' or 'Santa Claus'; it is plausible to say that these are only *pretended* to refer, they don't really. Since these terms are not meaningless, perhaps the associated descriptions must enter into an account of their meaning.

It is unclear that these represent a problem for Kripke. We can understand his position as that *normal* cases of proper names – what we called, in the box at the end of the 'Rigid designation' section, above, *genuine* proper names – don't fit the descriptivist theory, but that does not mean that none do. A more serious problem arises in cases such as that raised by Gareth Evans.⁷ It is said that the island of Madagascar was not originally so-called – that owing to a misapprehension on the part of European explorers and map-makers, the name which was originally the name of part of the African coast got transferred to the large island offshore. The practice of using it for the island gradually caught on. Unlike the 'Homer' case, one does not feel tempted to think that 'Madagascar' is a descriptive name. Surely it's a genuine proper name, subject to all the Kripkean arguments that show it to be so. Many subtly different variants of this sort of case are possible, but all have in common that the object denoted by the contemporary users of the name fails to match that of the original producers of the name. How precisely to handle such cases has proven a big topic; but presumably a solution has to do with the fact that the historical chain of communication gets *disrupted* in such cases, that at some point in history the users have begun passing along a piece of *misinformation*.

● LINGERING ISSUES FROM RUSSELL AND FREGE

Russell believed that ordinary proper names are not really proper names at all. They are abbreviations for definite descriptions, which according to his theory are not singular terms but quantifiers. The only logically proper names are names of sense-data (and perhaps the self). Logically proper names in his view have no descriptive content, contrary to Frege; they don't express a rule for determining a referent and they do not express a mode of presentation of the referent.

For these reasons, direct reference theory is Russellian because it sides with Russell and against Frege on the question of the descriptive content of proper names: they haven't any. It departs from Russell, however, in accepting ordinary proper names as genuine proper names – that is, as logically proper names.

It's true, Kripke's arguments against Frege and Russell's descriptive view of ordinary proper names are very powerful. What, then, of the epistemological reasons which led Frege to the theory of sense, and which led Russell to hold that only 'items of acquaintance' can be the referents of logically proper names? How is a direct reference theorist to explain the following sort of possibility?

Hesperus = Phosphorus.

John believes that Hesperus is a star.

John does not believe that Phosphorus is a star.

Cannot all three be true? For Kripke it now looks as if 'Hesperus' and 'Phosphorus' have the same meaning after all, and therefore the same cognitive content, in which case, surely, these three sentences cannot all be true. And what of names for the non-existent, such as 'Vulcan'? This is not a fictional name, and was intended for all the world as a normal proper name. But since it has no referent, it appears that Kripke's account gets the wrong answer that the name is meaningless.

There seem to be three possibilities. One might try to come up with some sort of grand unification theory, one that accounts for the cognitive dimension of language that exercised Frege and Russell but which accommodates the facts about reference and necessity that Kripke points out. Or one might be able to rebut Kripke's arguments. Or one might accept Kripke's arguments, but propose, *contra* Frege, that questions of *reference* should more sharply be distinguished from questions of cognition, which are matters for *epistemology*, the *philosophy of mind*, or *psychology*. We'll return to this in Chapter 11.

● FURTHER DISCUSSION: INTENSIONAL SEMANTICS

The 44th US president was the first black US president. The two definite descriptions happen to pick out the same individual, but they don't mean the same. This

is reflected in the fact that the two don't *necessarily* pick out the same individual: at other possible worlds, for example, the 42nd US president was the first black US president, not the 44th.

The only mammals that fly are bats (assuming that flying squirrels don't really *fly*, that they only *glide*). Therefore the predicates ' α is a bat' and ' α is a flying mammal' are co-extensive; the set of bats is exactly the same as the set of flying mammals. But the two predicates are not synonymous; their meanings differ. And this is reflected in the fact that although in the actual world the set of bats is the same as the set of flying mammals, there are other possible worlds where there are creatures that are the one but not the other.

Impressed by this correspondence, many philosophers have proposed that the meaning of linguistic expressions should in some sense be identified with this behaviour across possible worlds, their 'modal' features. We can speak of each expression as having an **intension**, which determines, at each possible world, its extension at that world. So the intension of 'the first black US president' determines Barack Obama as its extension in the actual world, but someone else, say Jesse Jackson, at another possible world, and others at others. The intension of ' α is a bat' determines the same extension as ' α is a flying mammal' in the actual world, but determines a different extension at other possible worlds, such as one with flying wombats.

Similarly, we can identify propositions with the intensions of whole sentences. We have been assuming since Chapter 1 that propositions are abstract objects whose structure is analogous to the sentences that express them. Instead, we now think of a proposition as a function from the sentence and a context of utterance to sets of worlds, namely the worlds at which the sentence is true (for that context of utterance). The intension of a sentence determines the truth-value of the sentence at each world. The truth-value of 'Barack Obama is the 44th US president' is the same as that of 'Voltaire wrote *Candide*'; but the equivalence is only an actual one; in other possible worlds the two have different truth-values. Meaning, we can say, is to be explored in terms of the modal behaviour of sentences and more generally of expressions. This fits with Frege's characterisation of the sense of a sentence as its truth-condition (and of an expression generally of its reference-determining condition): to say that the sense of a sentence is its truth-condition is to say that the intension of a sentence is the set of circumstances under which it is true.

This, or variants of this, have for the past 60 years been a popular approach for explicating meaning. It's very intuitive. One's grasp of the meaning of the word 'red' seems to involve knowledge not just of what counts as red in the actual world, but also what would be red in alternative situations, other possible worlds. The most reliable and familiar technique to convince yourself that one expression's meaning differs from another's is to engage in hypothetical reasoning: in the actual world, A and B stick together, but are there possible situations in which they fall apart? If so, then A and B have different meanings. But the other way around – that if A and B stick together in every possible world, then their cognitive meanings coincide – is not true. For if A and B are *themselves* necessary truths – say ' $2 + 2 = 4$ ' and 'Any kangaroo is a kangaroo' – we don't want to say that A and B thereby mean the same.

Intensional equivalence cannot entail cognitive equivalence, not just like that. Similarly, there are predicates that are cognitively inequivalent but necessarily have the same extension: ‘ α is a round square’ and ‘ α is an animal that lived before animals existed’ have the same extension – namely the empty or null set – but obviously don’t mean the same. For singular terms, the obvious cases are those definite descriptions built of the predicates just mentioned: ‘the round square’ and ‘the animal that lived before animals existed’; these are intensionally equivalent, since they denote nothing in every possible world, but they are not cognitively equivalent.

There are many, many responses and approaches to this set of problems, beginning with the Frege-based Carnap and Church in the late 1940s, through Montague, Kaplan, Lewis, Stalnaker and many others up to the present. An early approach was to appeal to the compositional structure of sentences. ‘The round square’ and ‘the animal that lived before animals existed’ necessarily have the same extension, but it’s not the case that they are built in the same way from parts which are themselves intensionality equivalent. That they should be is Carnap’s requirement of *intensional isomorphism*: synonymous sentences – sentences that are cognitively equivalent – must be intensionally isomorphic, that is, built in the same way from the same intensions.

But this and others like it are too complex to investigate in detail for this book – as well as there not being one that dominates the scene in the ways that the ideas of Frege and Russell have. Instead, one might suppose that getting the modal behaviour right is a *necessary* condition, if not a sufficient condition, of a successful semantic theory. A theory of meaning must not give wrong answers to questions of truth-conditions for sentences across possible worlds.

● HISTORICAL NOTES

Until the 1960s, many well-known philosophers – Hume and Ayer being conspicuous examples – accepted that the three distinctions – *a priori* versus *a posteriori* (or empirical), necessary versus contingent and analytic versus synthetic – aligned; moreover, the first two distinctions could be explained as, or reduced to, the third. There is no need for a queer faculty of *a priori* knowledge and there is no mystery of how truths can be necessary: all of that can be explained in terms of semantics – in terms of how we speak; what can look like metaphysically or epistemologically substantive matters are unsubstantive matters concerning language. As we saw in Chapter 4, the Logical Positivists, armed with the much more powerful logic of Frege and Russell, from the late 1920s through to the 1950s developed the view, seeking an end to metaphysics and new clarity to epistemology. Carnap’s *The Logical Structure of the World* (1967 [1928]), *The Logical Syntax of Language* (1937 [1934]), and ‘Empiricism, Semantics and Ontology’ (1950, in Carnap 1956), and Ayer’s *Language, Truth and Logic* (1936, revised edition 1946) were clear and influential statements of such views. Kripke published his paper ‘Identity and Necessity’ in 1971, and the book that made him and his views famous, *Naming and Necessity*, in 1980 (it was first published in 1972 in *Semantics of Natural Language*, eds D. Davidson

and G. Harman). Contra Ayer and Carnap, the three distinctions are actually quite separate and examples of propositions were given that are *necessary* but *empirical* and *a priori* but *contingent*. Kripke's views were enormously influential, not only on philosophers of language but on epistemologists, metaphysicians and philosophers of science. In general, the views were liberating: after years of nay-saying by Carnap and his ilk, philosophers could now openly advance metaphysical theories without the supporting proviso that the claims are implicit in our concepts, that they are analytic. Like many significant changes in intellectual history, however, Kripke was only the most visible of the revolutionaries; Peter Geach, Arthur Smullyan, Dagfinn Føllesdal, David Kaplan, Keith Donnellan as well as Hilary Putnam (see Chapter 8) were writing things in the 1960s very much within what was often called the 'new theory' of reference, and the idea is present in the work in modal logic of Ruth Barcan Marcus dating from 1947 in her characterisation of names as descriptionless 'tags'. The legend is that the key event was a meeting of the Boston Colloquium for Philosophy of Science in 1962, at which Barcan Marcus presented the results of her 1961 paper, 'Modalities and Intensional Languages'. In the audience were W. V. Quine, Alfred Tarski and the young Saul Kripke, who had already made a name for himself at age eighteen with a 1959 paper on modal logic. What happened that day, what precisely was understood by whom, and how much was retained by Kripke, remains controversial. What is relatively clear is that although Barcan Marcus discovered the idea, she did not pursue its philosophical implications nearly so far as Kripke did.

● CHAPTER SUMMARY

Possibility, in the sense we have in mind, is a metaphysical notion: to say that *p* is possible is to say that, even if in fact *p* is not the case, and even if it is known that *p* is not the case, it might have been the case that *p*. In this sense, that *the ancient Greeks developed nuclear power* is a possible but not actual truth. Possibility is inter-definable with necessity: *p* is possible if it is not necessary that not-*p*; *p* is necessary if it is not possible that not-*p*. At worst, the notion of a *possible world* is a very useful heuristic; it is a way that things could or might have been.

Frege and Russell can be thought of without too much violence as having agreed that for any ordinary proper name *N*, it means on the particular occasion of its use or understanding some definite description of the form *the F*. Kripke poses no fewer than four objections: (1) if *N* means the *F* then it is *necessary* that *N* is the *F*; typically that is not so; (2) if *N* means the *F* then it is *a priori* that *N* is the *F*; typically that is not so; (3) if *N* means the *F* then it is *analytic* that *N* is the *F*; typically that is not so; (4) one can understand sentences containing *N* without accepting *any* proposition of the form '*N* = the *F*' (where *F* is not itself a linguistic condition). One can refer to a thing by means of a name without possessing descriptive knowledge of the thing.

Kripke provides a *causal-historical* alternative to the descriptive theory of what determines reference: one uses a name *N*. What do I refer to by *N*? One acquired *N* from language-user *A*; and what did *A* refer to by *N*? *A* acquired *N* from *B*; and what

did B refer to by N? We follow the chain back in time until we reach a person who fixed its referent by either dubbing – saying ‘this object shall be named “N”’ or some equivalent – or by description, as in ‘I shall name the F “N”’. It is crucial that in the latter case, the description merely points to the object; it is not to be understood as synonymous with the name, and the name does not thereby acquire any conceptual content. Furthermore, the criteria for the reference of a name do not have to enter into one’s understanding of the name; there is no need to know how the chain links one back in time to a reference-fixing act.

Rigid designators are terms which always refer to the same object in every possible world (except those in which the object does not exist). Others are called non-rigid or flaccid designators; an example of the former is ‘the number of sides in a square’; of the latter, ‘the 42nd US president’. Ordinary proper names are rigid designators; thus ‘Charles Dodgson = Lewis Carroll’ is a necessary truth, whereas ‘Lewis Carroll = the author of *Alice in Wonderland*’ is true but not necessarily true. ‘Charles Dodgson might not have been Charles Dodgson’ is false, but ‘the author of *Alice in Wonderland* might have not been the author of *Alice in Wonderland*’ is true.

Intensional semantics is an attempt to explain meaning via the concept of possibility; such observations as that knowing a sentence’s meaning enables one to determine, for any possible world, whether it is true, encourages one to identify meaning with modal truth-conditions. One hang-up is that it is implausible to say that knowledge of its truth-value in every possible world is sufficient for knowledge of the meaning of a sentence, since, for example, not all necessary truths are synonymous with one another.

● STUDY QUESTIONS

- 1 Consider the suggestion on behalf of the descriptivist theory of proper names that there *is* a description that any user of the name N can supply which will refer to the right object: ‘the object that the linguistic community *actually* refers to by N’. Run through each of Kripke’s four objections to the descriptivist theory. Do they apply here? Does the suggestion involve a vicious circle?
- 2 Does a person not need *any* conceptual resources to refer to an object? It seems one must need them in order to do something more than just making noises. What, then, must a language-user understand or intend in order to refer by means of a proper name to an object?
- 3 Suppose Mrs. Smith goes to the hospital and gives birth to a boy. The squiggling thing is handed to her; she says ‘I dub him “Roger”’. He is taken away to be weighed, checked over and cleaned up. But while undergoing this procedure he gets mistakenly switched with another boy who was born at same moment, called ‘Sam’ by his mother. Nobody ever knows that there was a mistake. Brought to Mrs. Smith, the boy who was dubbed ‘Sam’ grows up happily, being called ‘Roger’. What is the boy’s name? Years later, if the mistake is detected, does the boy learn that he is not Roger, but Sam?
- 4 There is an iron bar held in Paris – called the ‘standard metre’ – that used to be regarded as the ultimate criterion for the length one metre. Thus one would

be forgiven for supposing that at that time, ‘the length of standard metre-bar = one metre’ was a necessary truth. But Kripke disagrees. Contrary to what Wittgenstein said about the matter, it was *a priori* but contingent that (at a certain moment) the length of the standard metre-bar is one metre, for the bar could have been longer or shorter. What is true is that (at a certain moment) the ‘length of the standard metre-bar is one metre’ served to fix the reference of ‘one metre’. Why would he say this? Is he right?

- 5 Consider ‘Snow is white’ and ‘Snow is white and all kangaroos are kangaroos’. Do these have the same intension? What can be said from the point of view of intensional semantics to reconstruct the strong intuition that they do not mean the same?

● PRIMARY READING

Kripke, S. (1980) *Naming and Necessity*.

● SECONDARY READING

For a recent tour of the Kripkean landscape see Stanley, J. ‘Names and Rigid Designation’, in Hale, B., Wright, C., and Miller, A. (eds) *A companion to the philosophy of language*, second edition. For some early incarnations of the ‘New Theory of Reference’, see *Reference and Modality*, edited by L. Linsky (1971); which contains pieces by Smullyan, Barcan Marcus and Føllesdal. For a kind of pushback by G. Evans see his ‘The Causal Theory of Names’ in his *Collected Papers* (1985), and *The Varieties of Reference* (1982); for a sophisticated and modern examination of both Evans’s view and Russell’s ‘Principle of Acquaintance’, see Hawthorne and Manley (2012) *The Reference Book*.

● NOTES

- 1 Some believe that the main idea behind Kripke’s view of proper names was first hatched in the pioneering work in modal logic by Ruth Barcan Marcus; see the ‘Historical notes’ at the end of this chapter.
- 2 It is possible to interpret ‘is necessary’ as a predicate of propositions, or indeed of sentences themselves. The approach runs into obstacles and is not widely taken up. See Skyrms 1978.
- 3 Not everyone thinks that Frege accepts this. But perhaps this does not matter so far as Kripke is concerned: Kripke’s attack is directed at the descriptive theory itself, whoever held it.
- 4 A complication: this is so only if no indexicals occur in the relevant terms. Kaplan says that a sentence like ‘I’m here’ is analytically true – the conventions of language are such that no utterance of it can ever be false – yet the propositions

expressed by it are contingent. For example, suppose one is at Wembley Stadium; thus an utterance of 'I'm here' expresses the proposition that one is at Wembley. But it is not a necessary truth that one is at Wembley, for one could have been somewhere else. We'll investigate further in Chapter 8.

- 5 If Mozart were never born, then would not 'Mozart = Mozart' fail to express a truth? Some think that 'Mozart = Mozart' would be true even if Mozart didn't exist. Or we could write 'if Mozart exists, then Mozart = Mozart'. Likewise we could write 'if the composer of *Don Giovanni* exists, then the composer of *Don Giovanni* = the composer of *Don Giovanni*'. But that would complicate the discussion unnecessarily.
- 6 To be more precise, we can say that a rigid designator designates the same object *in every world in which that object exists*. What Kaplan calls an *obstinately rigid designator* designates the same object in every world, period – in which case the object is a *necessary existent*.
- 7 For this and many other examples, see G. Evans, 'The Causal Theory of Names', in his *Collected Papers* pp. 1–24; for 'Madagascar' see p. 11.

8

• context-dependence, indexicality and natural kinds

● INDEXICALS AND DEMONSTRATIVES

Consider the following words:

I, me, myself	tomorrow
you, yourself	today
he, himself	yesterday
she, herself	this, that, it
him, her	his, hers
we, us, ourselves	now, then
they, them	here, there
yours, theirs	mine
those, these	soon, when

These are all indexicals (pronouns; called 'deictic' expressions in linguistics). Take the word 'I'. It is a singular term, but it would be wrong to say that the *word* 'I' has a referent; it is not like 'Rotterdam', always having the same referent on each occasion of use. Rather, each *utterance* of the word normally has a referent. Its referent is the speaker, the one who happens to be saying it.

Take the word 'today'. Each utterance of the word refers to a day, namely the day containing the moment at which it is uttered. The word 'here' refers to the place

at which it is uttered (although ‘here’ also has a *demonstrative* use; see below). And so on.

In general, we say that the reference of an indexical depends on the ‘context of utterance’. This is not so of an expression like ‘the sun’ or ‘Rotterdam’, which always refers to the same thing irrespective of the context of utterance.

Thus the linguistic meaning of an indexical is a *rule* (a *function*) that determines the referent of a token of that indexical at each possible context of utterance. For each possible context of utterance, the rule specifies what the referent of an utterance of the indexical would be at that context.

Tense represents indexicality with respect to time. When one says ‘It’s snowing’, one means that it’s snowing *now*. For the past, one says ‘It snowed’ or ‘It was snowing’; for the future one says ‘It will be snowing’, ‘It will snow’ or ‘It’s going to snow’.

Some indexicals are **demonstratives**. Take the word ‘that’. In its demonstrative use, if no such object is ‘salient’, then the speaker must, in order to refer to something, point or otherwise gesture towards the intended referent. If, in a forest, a speaker says ‘That’s a very beautiful tree’, but doesn’t point or otherwise indicate which tree they mean, then it seems they have not succeeded in expressing what they meant to say. In order to do so, they must point to or otherwise indicate the tree they mean. But the pointing gesture may be *implicit* rather than *explicit*. If we are walking through open country, and there is just one tree to be seen and it’s right there in front of us, then I needn’t point out which tree I mean when saying ‘That’s a very beautiful tree’. You’ll know which tree I mean.

The gesture needed to complete a demonstrative is called a **demonstration**. Indexicals such as ‘now’ that do not require a demonstration – whether tacit or overt – are called **pure indexicals**. The referent of an utterance of a pure indexical is determined simply by the context of utterance.

The context of utterance is roughly the situation in which an utterance is made or could be made. It includes, but is not restricted to:

(identity of) speaker, audience, time, place.

Unlike ‘I’, ‘now’ and the like, which are simple, many indexical expressions are **compound** – for example, ‘my horse’, ‘that big red house’, ‘the day before yesterday’ and so on.

Many words have both an indexical and a non-indexical use. Consider:

Every man who owns a donkey beats it.

Lady Ottoline Morrell met Ludwig Wittgenstein, but didn’t like him.

‘It’ and ‘him’ in these contexts are reflexive pronouns, devices of *cross-reference*. Someone using these words in these ways is not relying on the context of utterance to determine their referents. ‘It’ in the first example works like a bound variable ‘x’ in logical notation, and ‘him’ in the second case could without loss be replaced by

another occurrence of 'Ludwig Wittgenstein'. In the first case, there is no particular donkey being referred to at all; in the second, the referent of 'him' is determined by a so-called **anaphoric** connection to the occurrence of 'Ludwig Wittgenstein'.

In the following cases, however, those same words do function as indexicals:

(One surfer to another, as they watch an incoming wave) It's a big one!

(One prison camp guard to another, as a prisoner is seen escaping through the barbed wire) Shoot him!

The fact that, unless we are talking long-distance on the telephone or something like that, we do not normally say 'It's snowing *here*', but just 'It's snowing', shows that the types of indexicality, or, as we will sometimes say, the type of **context-relativity** or **context-sensitivity** of what we say, is often *implicit*. Frequently, it simply goes without saying what one is talking about; our intended audience will understand.

The use of quantifiers is especially striking in this regard, and is important for making sense of ordinary discourse. For it's rare that by 'every' we really mean 'everything whatever'. For example, if one says 'Everything's cooked' just before a dinner party, one doesn't mean that *Everything that exists in the universe is cooked* – or that *Everything that exists, has existed or will exist is cooked*. One merely means that the food for the upcoming dinner party is cooked. One *implicitly restricts the domain of quantification*, relying on the hearer's knowledge of the situation to interpret what one says. That is, one interprets 'Every x is –' as if it meant 'Every x which is F is –', where F is a predicate delimiting the range of the quantifier (in this case the place of F would be taken by 'food that one has actually prepared for the upcoming dinner party').

● PUTNAM ON NATURAL KIND TERMS AND ESSENCE

Putnam's famous 'Twin Earth' argument runs as follows (in 'The Meaning of "Meaning"', Putnam 1975). Water, as you know, is H₂O. Suppose there were a planet that is exactly like Earth in every respect, down to the smallest detail, except that the clear liquid in rivers and streams is some other chemical compound XYZ. For all ordinary purposes XYZ is indistinguishable from water; it behaves exactly as water does – it boils at the same temperature, sustains life and so on. Would that stuff *be* water or not? Putnam says not. So long as it is not H₂O (or H₃O, but ignore this), it is not water, no matter how exactly similar it is in other respects to water.

Why? According to Putnam, when we use a term such as 'water', we are using what is called a **natural kind** term. In fact, it functions almost exactly as a proper name does according to Kripke. Idealising somewhat, we define it – fix its reference – by pointing to some water and declaring: by 'water' I mean *that stuff* (just as the parents might fix the reference of 'Josef Haydn' by declaring 'By this name we mean this child'). Similarly, to fix the reference of 'Tiger' we point to an animal and decide that that word shall apply to all and only animals of the same species as *that animal*.

Natural kinds include substances like water and gold, and species like tigers and magpies. More generally, we introduce a natural kind term by pointing to a sample of the kind and intending that the word shall denote whatever is the same kind as the sample. The role of 'same natural kind' with respect to these sorts of predicate is exactly analogous to that of the 'same object' with respect to proper names.

It follows that a descriptive definition of such a term as 'gold' cannot be correct. One might try to define it as a 'shiny, malleable yellow metal'; we may suppose this description would be *true of* gold and only of gold, but it does not give a *synonym* for 'gold'; it does not actually give the meaning. For consider:

Gold = the shiny, malleable yellow metal that is widely used for jewellery.

This is undoubtedly true, but it is only contingently true. It is only contingent that it is used for jewellery. And in other possible worlds there are other shiny, malleable yellow metals (perhaps there is an actual example in iron pyrite). And one may compare the arguments of the last chapter concerning proper names: just as it is not a necessary truth that gold satisfy the above definition, it is not a necessary truth that Mozart was the composer of *Don Giovanni*.

What is it for something to be the *same kind* as another thing? In the case of elements, such as gold, science says that one thing is the same kind as another if and only if they have the same atomic number (the number of protons in the nucleus of the atom; for gold it's seventy-nine). In the case of chemical compounds such as water, it is the composition of molecules; in the case of water, each molecule comprises one atom of oxygen and two of hydrogen: H_2O . In the case of species, it is less clear, but one account might be: being a member of a certain genealogical type; another might be having a certain DNA configuration; another might be capacity for mating to produce fertile offspring (I say 'might be' because the matter is contentious, a live topic in the Philosophy of Biology).

Natural kind terms are not descriptive like 'muddy' or 'vermin', but certain descriptions do have a privileged status with respect to natural kinds. If we are pointing to some water, then necessarily, something is *that stuff* if and only if it is H_2O . That is because *having the same molecular constitution* is the criterion of identity for compound substances. Therefore:

Necessarily, water = H_2O .

So it seems we can say: H_2O is the *essence* of water; H_2O is just what water is.¹ But this is not something we know *a priori*, and it is not the case that 'water' and ' H_2O ' are synonymous. A person who does not know that water is H_2O lacks scientific knowledge, but does not thereby lack linguistic knowledge: for they too use the word in sentences and can point to water and say that 'water' means that stuff.

Thus, contrary to what many philosophers had thought (Hume, Kant), *some necessary truths are empirical (a posteriori)*. This opens the door to some weird possibilities. For example, it might seem analytic or necessary that cats are animals. But in the fanciful but conceivable possibility that the things we have called 'cats' are and

always have been robots controlled by aliens, rather than there being no cats, 'Cats are animals' would be false.

An important warning: a common mistake is to overreact to Putnam (and Kripke) and to think that what has been shown is that *all* general terms are natural kind terms, that all terms pick out an essence or that all terms refer directly to their referents or are rigid designators. That's not so. Not even all nouns are directly referential or are rigid designators. The concept of a *hill*, or *mud*, for example, cannot be defined in terms of the nature of a substance or in terms of the nature of the things that actually realise the role marked out by the word. Unlike 'water' or 'tiger', it is not conceivable that hills or mud could turn out to have hidden natures (let alone natures quite other than what we always took them to be). Unlike 'water' or 'tiger' – like 'watery' or 'tiger-shaped' – the terms express concepts of the roles themselves, not the realisers. Similarly, adjectives (like 'stupid', 'pretty'), verbs ('shake', 'thwart') or adverbs ('busily', 'centrally') are not plausibly construed as natural kind terms. Much of language comprises descriptive words that are not directly referential or rigid (a good question concerns terms for *artefacts*; is the concept, say, of a *book*, a concept of a certain artefactual essence – are there *artefactual kinds* – or is it a descriptive concept?).

There are similarities between Putnam's account and John Locke's old discussion of real and nominal essence. For Locke, the nominal essence of a kind is the rule we use to pick out instances of the kind, normally comprising observable properties. So the nominal essence of gold might be 'shiny, yellow, malleable metal'. The real essence of a kind is its underlying nature; Locke was pessimistic about whether we could know real natures, but surely he'd be glad to learn that he's been proven wrong.

Putnam's point about natural kinds can be summed up by saying that terms for natural kinds typically function as *indexicals* such as 'this', and are not equivalent to any context-free description such as 'big orange stripy carnivore'.² What reference gets fixed by introducing such a term depends on the nature of the animal or whatever it is that is actually pointed to when the term is introduced.

In fact, as was suggested above, that is very much what is going on with Kripke's explanation of proper names: proper names share many features with indexicals, and among them is that the bearer is determined by the identity of the object pointed to at the starting point of the causal chain when someone says 'Let's call him "Josef"' etc. Just as in the case of ordinary proper names, one can introduce a natural kind term by means of a reference-fixing description: one might say 'by "water" I mean the stuff that *actually* flows in rivers and falls as rain', without thereby deeming 'water' and 'the stuff that actually flows in rivers and falls as rain' synonymous.

● IS MEANING IN THE HEAD?

From a Fregean point of view, it would be natural to suppose that the proposition one expresses by means of a sentence such as 'Daphnia live in water' is determined by one's internal psychological state – ultimately by the state of one's brain. It follows from Putnam's indexical view that this is false. If Twin Earth is really just like Earth, then your doppelganger lives there: your exact duplicate except that they have XYZ

in the immediate environment whereas you have H_2O . If so, then when you and your duplicate say ‘Daphnia live in water’, you mean different things. Since you refer to H_2O , and your duplicate to XYZ, the truth-condition of what you say is that Daphnia live in H_2O , and that of what your duplicate says is that Daphnia live in XYZ. Since you and your duplicate must be in the same brain-states (assume that the identity of brain-states does not depend on whether it’s water or XYZ that’s inside the body), it follows that the meaning of what you say is not determined by your brain-states. It depends also on the environment, on the physical context in which you live. What we know as ‘water’ is not the same as what those on Twin Earth know as ‘water’. Thus if you were presented with a glass of XYZ and said ‘Ah, water!’, you’d speak falsely, even though Twin Earthlings call it ‘water’. The words ‘water’ on Earth and ‘water’ on Twin Earth would be homonyms – words that look and sound the same – but not, strictly speaking, synonyms. Such is the basic argument for *semantic externalism*: you and your duplicate refer to different things despite there not being a difference in what concepts you grasp, or there not being a physiological difference between you. As it’s often put, semantics does not *supervene* on physiology or, more generally, on one’s internal states. This should be no more surprising than the fact that you and your doppelganger express different propositions by ‘I love my mother’.

Another striking piece of data Putnam cites for the externalist viewpoint is this. Suppose you know that the words ‘elm’ and ‘beech’ are terms for deciduous trees, but that you know no more about them. You know nothing about the one that you don’t know about the other. Still, if you point to a beech and say ‘That’s an elm!’, what you say would be false; if you’d said ‘That’s a beech!’, what you would have said would have been true. What is going on is the operation of what Putnam terms the **division of linguistic labour**. The semantics of what you say is determined not merely by what transpires in your head, but also by the capacities of others to make the necessary distinctions. To take another example, suppose in the year 1800 you find some yellowish metal. ‘Gold!’, you declare. Your sister is more sceptical. You take it down to old Mr. Burn’s jeweller’s shop, trusting him to be able to tell definitively by administering accepted tests (the modern definition in terms of atomic numbers was in the future). In fact, it wasn’t gold; your initial declaration was false, and your sister was right to demur. The division of linguistic labour involved the jeweller’s expertise – a certain deference, not to Mr. Burns, but to the practice of counting the tests as definitive.

● THE ACTUAL WORLD AS A CONTEXT

Putnam seems to show that the surrounding environment shapes the meaning of words like ‘water’ more radically than we might have thought, and what goes on in our minds rather less so. It boils down to potential differences in the *context of utterance* of, say, ‘This is water’. Let us say more to nail down exactly what this means and to draw out some repercussions.

Suppose Jones exists in a possible world W . His dog is Fido. Jones says ‘My dog has fleas’ at exactly 1.00 p.m. What he says is true just in case Fido has fleas at 1.00 p.m. If Jones utters the very same sentence again at 8:00 p.m., then what he says is true

just in case Fido has fleas at 8.00 p.m., not 1.00 p.m. *What he says* depends on the context of utterance: the proposition he expresses on the second occasion is different from the one he expresses on the first, because on each occasion the present tense of the verb 'to have' indicates that the dog has fleas *now* – i.e. at the time of utterance, which changes between one utterance and the other.

Now consider Jones in an alternative possible world W^* , in which Jones has a different dog, Spot. He says 'My dog has fleas' as before. Clearly, what he says is true just in case Spot has fleas at the time of utterance.

Now return to the first world W , where Jones' dog is Fido. Suppose now that Fido doesn't have fleas, but Jones says 'My dog might have had fleas'. What he says could be explained in either of two ways:

- (1) In some possible world, Jones has a dog, and it has fleas.
- (2) In some possible world, Fido has fleas.

It seems clear that (2) is correct. When Jones says 'My dog might have had fleas', he is talking about his *actual dog*, the dog he has at W . He means that *it* might have had fleas: that there is some possible world in which that very dog, Fido, has fleas. He does not mean that in some world he has some-dog-or-other that has fleas.

Now back over to W^* . Assume that Spot does not have fleas in W^* . If Jones had said, in *that* world 'My dog might have had fleas', then he is talking about Spot, not Fido. Which dog Jones is talking about using 'my dog' depends on which world he is in. So we have to add the *world* of utterance to our list of indices that make up the *context* of utterance:

Contextual indices: time, place, speaker, audience, and *world*.

Of course, no sentences are actually uttered at non-actual possible worlds. But sentences are uttered at such worlds (just as other things happen: apples fall from trees, wolves howl and so on). So the context of utterance is not *necessarily* the actual world, even though the context of every actual utterance is the actual world.

'My dog' is both an indexical and a rigid designator. It is indexical since it means 'the dog that belongs to *me*'. Its being rigid was why, in evaluating 'my dog has fleas' at other worlds, we were asking the question about the *same dog* in each world.

This is not to deny that there are also non-rigid uses of phrases like 'my dog'. For example, we can say 'I wish he were not my dog', without thereby wishing he were not the dog that he is; such uses are equivalent to non-rigid descriptions such as 'the dog that belongs to me'. Whether or not such an expression is being used rigidly sometimes depends on the intentions of the speaker. Thus assume that my dog is a beagle, and consider:

If I had chosen a different dog, then my dog would have been a French Bulldog.

If I had chosen a different dog, then my dog would have belonged to someone else.

In the first sentence, 'my dog' is used non-rigidly (I am not saying that my beagle might have been a French Bulldog!); in the second sentence, it is being used rigidly (I am not saying that a dog might have been both mine and someone else's).

● TWO-DIMENSIONALISM: CONTEXT OF UTTERANCE VERSUS CIRCUMSTANCE OF EVALUATION

We acknowledged at the end of Chapter 7 an itch you may have felt during the discussion of Kripke, and you may be feeling it again now with Putnam. Suppose we grant that water is indeed necessarily H₂O. Still, you might think, surely in some sense water *could have turned out* not to be H₂O, but XYZ. That is surely imaginable, surely *conceivable* in some sense. Indeed, that is precisely what one is being asked to imagine when imagining Twin-Earth! If the XYZ-world *had been actual*, then water *would have been* XYZ – even though, in actual fact, water is necessarily H₂O. What is now known broadly as Two-Dimensional Semantics addresses this.

If all goes well, an utterance of a declarative sentence expresses a proposition: something that can be true or false. The proposition is the meaning of the words spoken in the context; it is *what is said*. We have just said that among the contextual parameters that determine what is said is the *world* in which the utterance takes place (along with time, place, identity of speaker and so on). We have also said, however, that we can take any proposition, and evaluate it *at*, or *with respect to*, a possible world; we can ask whether it is *true* at that world. The proposition has a *truth-value* at each world.

Thus, following David Kaplan, we are led to distinguish sharply between context of utterance and **circumstance of evaluation**:

Sentence-meaning + context of utterance \Rightarrow proposition

Proposition + circumstance of evaluation \Rightarrow truth-value

A *sentence-meaning* is a linguistic rule for determining the proposition expressed by a given sentence for each possible context of utterance: in Kaplan's terms, the **character** of a sentence determines, with respect to a given context of utterance, what **content** is expressed, what proposition is expressed. We say more generally that the character of a sentence is determined by the characters of the expressions it comprises, including its indexical expressions. For instance, the character of the indexical pronoun 'here' is wherever the utterer happens to be; its content, for a given utterer at location L, is L.³

We may then ask whether that proposition (that content) is true at a given possible world. When we say that a proposition is true, we normally mean truth at the actual world; but we may also ask of the proposition expressed whether it is true with respect to other worlds. Truth, in this way of picturing things, is not absolute, not simply a property of propositions: it is a relation between a proposition and a world.

Certain consequences of this scheme are striking. Consider Jones' utterance of the sentence:

I am here now.

No actual utterance of this sentence could ever be false: the proposition expressed by any actual utterance of that sentence is always true with respect to the actual world.

This might lead one to suppose that since it could not be false: 'I am here now' expresses a necessary truth. That would be a mistake. For Jones says it at a certain time, in a certain place. What he says is that *Jones is in that place at that time*. Is what he says a necessary truth? No, of course not. Jones *could* have been in a different place at that time. In other possible worlds, he is not in that place at that time. So the proposition is false at such worlds.

We might, however, want to say that 'I am here now' is a kind of *analytic truth*: for linguistic rules – conventions of language – ensure the truth of what it expresses on each occasion of its use. Yet what it expresses on each occasion is a contingent truth, not a necessary truth. This flies in the face of the traditional philosophical view that analytic truths are necessary truths, but it appears that they just aren't.

Suppose I say:

I do not exist.

Although the proposition I express is false, there *are* circumstances with respect to which the proposition is true, namely worlds in which I don't exist. So it is not a necessary falsehood, even though every utterance of the sentence is true. This perhaps explains the wonder of Descartes' *cogito*: according to linguistic rules, every utterance of 'I exist' expresses a proposition that is true at the world of utterance; but even so, that proposition may be false at other worlds. So the sentence 'I exist' is an analytic truth, but the propositions it potentially expresses are not necessary truths.

In short, the context of utterance determines the proposition expressed by the utterance; the circumstance of evaluation determines the truth-value of the proposition at a given world. The context of utterance is a situation with respect to which we are asking: what is the proposition that is expressed by this sentence? The circumstance of evaluation is a possible world with respect to which we are asking: is this proposition true with respect to this world?

The picture, then, is as follows. We can think of two kinds, or *two dimensions*, of valuation that determine the modal characteristics of any assertion: context of utterance and circumstance of evaluation. What we mean by the contingency of 'water = H₂O' is contingency along the first dimension, as we consider alternative contexts of utterance of the sentence. What we mean by the necessity of 'water = H₂O' is necessity along the second dimension, as we consider alternative circumstances of evaluation. Sometimes the first dimension is termed *conceivability*; thus it is conceivable that water ≠ H₂O, but necessary that water = H₂O.

Table 8.1 Two-dimensional diagram for ‘water = H₂O’

	CE1: watery stuff = H ₂ O	CE2: watery stuff = XYZ
CU1: watery stuff = H ₂ O	‘water = H ₂ O’ is true ‘water = XYZ’ is false	‘water = H ₂ O’ is true ‘water = XYZ’ is false
CU2: watery stuff = XYZ	‘water = H ₂ O’ is false ‘water = XYZ’ is true	‘water = H ₂ O’ is false ‘water = XYZ’ is true

Many people find the likes of Table 8.1 helpful. Let ‘watery stuff’ be a non-rigid designator for whatever at a given world satisfies the stereotype that is actually satisfied by H₂O – a clear liquid that boils at 100 degrees Celsius, etc. Each horizontal row represents a possible context of utterance (or ‘world considered as actual’, as it is sometimes put), abbreviated ‘*CU1*’ or ‘*CU2*’. Each vertical column represents a possible circumstance of evaluation (a possible world), abbreviated ‘*CE1*’ or ‘*CE2*’.

The top two rows – *CE1/CU1* and *CE2/CU1* – are what Kripke and Putnam were primarily concerned with. Are there worlds with respect to which the proposition actually expressed by ‘Water = H₂O’ is false? The answer to that *metaphysical* question is no. The left column – *CE1/CU1* and *CE1/CU2* – indicates the imaginability or conceivability of alternative scenarios. Is it conceivable that ‘water = H₂O’ could have turned out false? The answer to that *epistemological* question is yes.

Especially interesting are *CE1/CU1* and *CE2/CU2*: the situations represented by the left-to-right, top-to-bottom diagonal. The first represents simply the actual non-modal proposition expressed by ‘Water = H₂O’; the second represents what *would have been* expressed by the non-modal sentence expressed by ‘Water = H₂O’ *had the XYZ-world been actual*. In general, a location on the diagonal represents the non-modal proposition that would have been expressed by a sentence had that situation been actual. Had the XYZ-world been actual, then in that world it would have been *necessarily* false to say ‘Water = H₂O’; it would have been false to say ‘Water = H₂O’ with respect to a world in which the watery stuff = H₂O. That is what is shown by *CE1/CU2*.

● FURTHER DISCUSSION: RIGID DESIGNATION AGAIN

Consider the following sentences:

- (3) The world record holder in the men’s 100 metres is Jamaican.
- (4) In 1990, the world record holder in the men’s 100 metres was American.

At the actual context of utterance, 6 October 2023, (3) is true: the definite description ‘the world record holder in the 100 metres’ denotes the *present* (the time when I was writing this) record holder, Usain Bolt of Jamaica. But the description as it occurs in (4) does not pick out Bolt; (4) does not say, of Bolt, that *he* was American

in 1990 (which would be false). The description picks out the record holder in 1990, namely Carl Lewis, and says truly of him that *he* was American. Compare:

(5) In 1990, the present world record holder in the 100 metres was American.

(5) says falsely of Bolt that he was American in 1990. If someone were to utter (5) with the intention of expressing the proposition expressed by (4), they would be misusing the word 'present'.

Such a definite description as 'the world record holder in the men's 100 metres' is normally used in a way that is called *temporally* non-rigid. This means that when we talk about different times, the object it denotes depends on what object satisfies the description at *that* time. So the description picks out Bolt in (3) but Lewis in (4).

Contrast this with 'the *present* world record holder in the men's 100 metres'. This description is *temporally rigid*. The denotation of the description is fixed by the context of utterance (namely the time), irrespective of the time being talked *about* – irrespective, that is, of the presence of such phrases as 'next year ...' or 'in 1990 ...' That is why (4), as I utter it now, is true, but (5) is false. The truth-value of (4) depends only on the nationality of whoever was the record holder in 1990.

Just as we can distinguish between *temporally* rigid and non-rigid terms, we can distinguish between *modally* rigid and non-rigid terms. Consider a variant on a previous example:

(6) That dog has fleas.

Imagine someone saying this of Fido – i.e. Fido is the dog demonstrated (pointed at) by the speaker. Now imagine someone pointing at Fido and saying:

(7) It might have been the case that that dog has fleas.

(7) is true just in case in some possible world, the proposition expressed in (7) by the clause 'That that dog has fleas' is true. Suppose we are checking a world to see if that proposition is true at that world. Which dog do we check for fleas? Clearly, we check *Fido*, the dog actually demonstrated in the context of the utterance of (7). No *other* dog is relevant.

The expression 'that dog', as it occurs in (6) and (7), is a modally rigid designator. What this means is that the context of utterance determines the referent of the expression, and even when we are talking about counterfactual circumstances, non-actual possible worlds, the referent remains the same; we keep talking about Fido, the dog determined by the context of utterance.

Compare:

(8) The winner of best-in-show at Crufts in 1966 had white hair.

(9) It might have been the case that the winner of best-in-show at Crufts in 1966 had black hair.

Both are true. But unlike ‘that dog’ as encountered in (6) and (7), ‘the Crufts winner of best-in-show in 1966’ is a non-rigid designator. The winning dog in 1966 was in fact Oakington Puckhill Amber Sunblush (OPAS), a white toy poodle. What makes (9) true is that, irrespective of that, in some possible world a dog with black hair wins the title – a different dog, say a giant schnauzer with black hair, might have won Crufts in 1966. Now compare:

- (10) It might have been the case that the actual winner of best-in-show at Crufts in 1966 had black hair.

Unlike (9), (10) can be read as being about OPAS: it instructs one to find the dog that actually won Crufts in 1966, and to find a world in which *that* dog is black. So it appears the word ‘actual’ does with respect to modality what ‘present’ does with respect to temporality: it tells us to use the actual context of utterance (time, world) to find the referent; it rigidifies what might otherwise be a non-rigid singular term.

● THE INDISPENSABILITY OF INDEXICALS

It is tempting to suppose that indexicals, in theory, are expendable. In principle we could get by with only singular terms that are *purely conceptual* – their meanings composed entirely of descriptive concepts so that the way in which the singular term acquires its referent does not depend on the context of utterance. Such a picture, however, seems impossible to sustain.

First, consider the case of a simple, pure indexical such as ‘here’. Suppose, standing somewhere in the New Forest, I say ‘There are no mushrooms growing here’. It is perfectly clear that I succeed in referring to the place where I’m standing, even if I have no other way to refer to it: even if I cannot *individuate* it, *pick it out*, by describing it using only general concepts or without exploiting the context of utterance. I might even say this when I am *lost*, so I have no way at all of describing where I am, except perhaps ‘somewhere in the New Forest’, which would fall far short of picking out the much smaller region I designate by saying ‘here’.

Second, there is a more systematic philosophical point. Many terms we use for times and places seem non-indexical such as ‘the solar system’, ‘California’, ‘4.00 p.m.’, ‘12 January 1986’. But there are a great many stars with orbiting planets; when we speak of ‘the’ solar system, what we really mean is *this* solar system. More generally, think of what is involved in locating things on a map. If I point to a certain location on the map as being the place we are to go, does that tell you what direction to go? No. It reveals where it is *only if you know where you are on the map*.

Similarly with times. Suppose you visit the planet Zog and are told that such-and-such happened or is to happen in the year 34909 in the Zog calendar. This doesn’t tell you when it happened unless you know what year it is *now* in the Zog calendar (you’d also need the criterion for a Zog-year). The situation is the same with our own calendar, but we are so used to it that it isn’t quite as obvious. If you are told that the date of a certain event is 12 January 1986, what you are immediately told

is that the date is exactly 1985 years and 12 days after the (supposed) birth of Jesus Christ. But that doesn't tell you *when* the event happened; strictly speaking, it doesn't even tell you whether the event is in the past, the present or the future. What you need to know is what date it is *now*. Thus what being told the date of a past event really tells you is that it took place a certain number of days before *now*, i.e. the present. So it seems you can't make use of these supposedly non-indexical descriptions of things without linking them to indexical ones. You require indexical anchor points to make use of language.

An even more striking case concerns first-person pronouns, as pointed out by John Perry. Suppose you are told that a certain person is going to be tortured. The person is described as 'the one person who is X, Y and Z'. You don't know who it is that is uniquely identified as being X, Y and Z. Your worry, naturally, is this: is the person who is X, Y and Z *me*? Your attitude towards the upcoming torture depends on this; no further descriptions of the person to be tortured, no descriptions of the form 'the person who is such-and-such', will tell you what you want to know unless you can determine whether or not the person who is such-and-such is *you*. The information conveyed – *the cognitive import encoded* – is not exhausted by the concepts involved in the description, but requires also the index indicated by the term employed.

Indexicals are thus essential to our understanding of language and ability to refer to things. Context-dependence is not a shortcoming or fault of language. Indeed, as Russell stressed, these are expressions of cognitive states such that one is apparently immune to misidentification: one always knows one is here, now, and oneself – even if one cannot relate the here and now to other systems of conveying spatial and temporal matters.

● INDEXICALS AND FREGEAN SENSE

Frege was well aware of indexicals and tried to incorporate them into his sense-reference theory. Speaking of the present tense (which implicitly contains what is expressed by *now*), he wrote:

If a time-indication is conveyed by the present tense one must know when the sentence was uttered in order to grasp the thought correctly. Therefore the time of utterance is part of the expression of the thought ... The case is the same with words like 'here' and 'there'. In all such cases the mere wording, as can be preserved in writing, is not the complete expression of the thought; the knowledge of certain conditions accompanying the utterance, which are used as means of expressing the thought, is needed for us to grasp the thought correctly. Pointing the finger, hand gestures, glances may belong here too.

('Thoughts', in Frege 1997, p. 332)

Remember that a sense can be understood as either a mode of presentation of the referent or a rule that picks out the referent. Frege's idea here seems to be that the linguistic meaning of an indexical is a rule that determines the sense expressed on

each possible occasion of its use; the sense thereby expressed determines the object. So the mode of presentation involves the context of use. The rule for determining an object by an utterance of, say, 'the present temperature' is something like 'the temperature at time *T* at location *L*', where *T* and *L* are the time and location of the utterance.

Frege also says, indeed in the very same passage: 'If someone wants to say today what he expressed yesterday using the word "today", he will replace the word with "yesterday"' (ibid.).

But Frege is skirting on trouble here. The trouble is that according to the longer quotation above, successive utterances of 'today' and 'yesterday', even if they do refer to the same day, present it in different ways. The 'conditions accompanying the utterance', which are 'part of the expression of the thought', are different. In that case 'today is Tuesday' spoken on Tuesday must differ in sense from 'yesterday was Tuesday' spoken on Wednesday. So once yesterday has passed, we *can't* express today what we expressed yesterday.

Frege's remark about 'today' and 'yesterday' is probably the wiser one, for the picture he sketches in the longer passage above is too restrictive. A bit later, in connection with the first-person pronouns 'I' and 'me', he says 'Everyone is presented to himself in a special and primitive way': the thought one expresses by using 'I' can only be grasped by oneself; for that thought, the listener must substitute a thought expressed using 'he' or some such. If so, then surely something analogous goes for 'now' and 'here': more than one person can be presented with a given moment and grasp it as *now*, but once that moment has passed, that moment cannot be presented *in that way* again. Similarly with 'here': only actually being in a place, can one understand it as *here*.

If that is so, then the thoughts expressed by utterances of 'I am hungry', 'It is raining here', 'It is raining now' and the like become unavailable if one is not the speaker, if one is not present or if the moment has passed. Yet this hardly seems to matter, as Frege observes in his remark about 'yesterday' and 'today': so long as we know we are referring to the right object, we don't care about the mode of presentation. If so, then the notion of sense does not seem quite so relevant to an account of communication, at least where indexicals are concerned. We'll return to this in Chapter 11.

● HISTORICAL NOTES

Elements of intensional semantics are found in Frege and Russell, and the general idea goes back much further. Modern attempts to carry out the idea began perhaps with Carnap's (1942) *Introduction to Semantics* in which he employs the notion of an '*L*-state' with respect to a semantical system as doing duty for possible worlds (an *L*-state is roughly an interpretation of the formal language). The theory of indexicals also had its precursors in the work of Frege, and more substantively in Russell's theory of what he called 'egocentric particulars' in his *Inquiry into Meaning and Truth* of 1940 (published in 1950). Until Kaplan's work officially appeared in 1989

in the form of his ‘Demonstratives’, various people had contributed to the topic, including Tyler Burge, Hector-Neri Castañeda, Gareth Evans, David Lewis, Barbara Partee and Frank Vlach. But with the appearance of Kaplan’s work, activity on the topic on the philosophical side exploded (actually manuscripts of ‘Demonstratives’ circulated much earlier than 1989). Hilary Putnam (1926–2016) made his name in various areas of philosophy – not only in the theory of reference as represented here, but also in the philosophy of mind, where he formulated the doctrines of functionalism and multiple realisability; metaphysics and philosophy of science, in which he formulated his model-theoretic arguments for internal realism; and in many other areas. Besides his work on semantics, John Perry has been enormously influential in metaphysics and the philosophy of mind. His main works specifically about indexicals are ‘Frege on Demonstratives’ (1977) and ‘The Problem of the Essential Indexical’ (1979).

● CHAPTER SUMMARY

Indexicals are referentially context-dependent expressions such as ‘now’ and ‘here’. Each has a character that determines, for each context of utterance, a referent, in such a way that the referent can shift with context. Some indexicals are demonstratives, requiring a demonstration, which may be implicit, to secure a referent.

The direct reference view of proper names extends very naturally to what are called natural kind terms, such as ‘water’ and ‘tiger’. If something is water, then it necessarily is water. But no true description in terms of observable or surface properties necessarily holds of water. However, a description in terms of what does distinguish water from other substances – H_2O – does refer to the same substance with respect to any possible world. Thus ‘water = H_2O ’ is a necessary truth, despite its being *a posteriori* and despite its not being analytic. Thus substances do have essences.

Many uses of indexical referring expressions refer rigidly. If one points at some water, and says ‘that is water’, one utters a sentence that expresses, on that occasion of its use, a necessary truth. Indeed, the phenomenon of rigidity is at its clearest in such cases; one feels that it is *that stuff*, unmediated by concepts, that one refers to.

It is entirely possible that there should be pairs of possible situations such that one cannot tell which one is in ‘from the inside’, but which differ in that one contains, for example, H_2O , whereas the other contains an undetectably distinct substance XYZ. A linguistic community that has always been in the first situation will refer to H_2O by the sound or inscription ‘water’, whereas a linguistic community that has always been in the second situation will refer to XYZ by the same sound or inscription. These are two words – homonyms – that look and sound the same. These considerations support the thesis of semantic externalism: the reference of at least some words is determined by the nature of the surroundings, not by internal conceptual resources available to language-users.

The context of utterance includes, in addition to the time and place of the utterance, the possible world in which the utterance takes place. We’re familiar with the idea

of asking whether, irrespective of its actual truth-value, a proposition would be true at alternative possible worlds. Thus we can distinguish between the context of utterance and circumstance of evaluation: the character of an expression – its linguistic meaning – plus the context of utterance yields the content expressed by an utterance of the expression; in turn, the content plus the circumstance of evaluation yields the truth-value, extension or referent of the utterance. The difference between character and content is obvious in the case of explicit indexicals, but interesting also in the case of natural kind terms such as ‘water’, which Putnam suggests contain an implicit indexical component. Given the actual world as context, ‘water’ has the content H_2O – it tracks H_2O across possible worlds (as circumstances of evaluation); but in other contexts, it tracks for example XYZ. This ‘two-dimensionalism’ captures the idea that despite its being (actually) necessary that water = H_2O , things might have turned out that water is XYZ, indeed that water is necessarily XYZ.

The modal issues have their ready temporal analogues; in particular the distinction between modal rigidity and modal non-rigidity is analogous to that between temporal rigidity and temporal non-rigidity. For example, ‘the present F’ is the temporal analogue of ‘the actual F’.

It does not appear possible to regard indexicality as a mere convenience. It appears that any map or calendar, for example, is useless unless one has an indexical anchor – similarly and strikingly, descriptions of the doings of human beings can never have the same force without the information that one of the people is *you*. To find that ‘the F is G’ is not intrinsically meaningful in the way that ‘I am G’ is.

Indexicality puts pressure on Frege’s ideal of communication, according to which perfect communication involves you and I grasping the *same sense* (the same thought or proposition). For example, the mode of presentation of a time signalled by ‘now’ passes instantaneously; it cannot later be presented in the same way. But that hardly seems to detract from one’s ability to communicate at a later time about the time indicated.

● STUDY QUESTIONS

- 1 Suppose we accept Kripke’s account of proper names. Suppose Brown says: ‘I am not dead’; does this mean the same as ‘Brown is not dead’?
- 2 Can the distinction between (9) and (10) of ‘Further discussion: rigid designation again’ be displayed with the scope of definite descriptions instead of the word ‘actual’? For example: ‘it might have been the case that the F is G’ versus ‘the F is such that it might have been the case that it is G’? Use examples.
- 3 Especially once we’ve observed that *tense* is a kind of indexical, should we conclude that all actual sentences contain indexicals?
- 4 To what extent should we be semantical externalists? Assuming we accept the Kripke–Putnam line on ‘water’ etc., should we also be externalists about the concept of a *table*? What about the concept of *food*?
- 5 Kripke (Lecture III of *Naming and Necessity*) has what is essentially a two-dimensional argument against physicalism in the philosophy of mind. Suppose we

grant that ‘water’ refers to the underlying nature of the stuff: a stuff which only contingently satisfies a description of its superficial properties. When we think of other worlds in which other substances with different natures satisfy that description, we are not actually thinking of water. You might think that ‘pain’ is similar: it refers to a state of the nervous system (assume that this is C-fibre stimulation), which is only contingently hurtful. Something other than C-fibre stimulation in other worlds which feels hurtful would nevertheless not be pain. But that cannot be right, Kripke says. If something feels like pain – is hurtful – then it is pain – regardless of what underlies it. So it is not possible to ‘reduce’ pain to any particular state of the nerves (one can imagine a similar argument for the irreducibility of *consciousness*). Is Kripke right?

- 6 The water in the Mississippi River is, alas, very polluted. Especially towards the delta, it contains, in addition to H₂O, many other substances too foul to mention. In fact, its proportion of H₂O as it passes through Mississippi is lower than that in a weak cup of tea. But despite this, the tea is tea, not water, whereas the polluted river-water is still water. Does this present a problem for Putnam? (Borrowed from Chomsky 2000, pp. 127–8).

● PRIMARY READING

Kaplan, D. (1989) ‘Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals’, in *Themes from Kaplan*, pp. 481–564.

Perry, J. (1979) ‘The Problem of the Essential Indexical’.

Putnam, H. (1975) ‘The Meaning of “Meaning”’.

● SECONDARY READING

A helpful recent survey of indexicality, including both remarks about its history and an argument for where things stand, is Manuel García-Carpintero’s 2017 piece ‘The philosophical significance of the De Se’ (*Inquiry* 60/3: 253–276). See the Chapter 14 section on context-relativity (including the reading notes) for more. The literature on externalism vs internalism is vast, but enormously influential at least from the standpoint of philosophy of mind/cognitive science is Jerry Fodor’s *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge, MA: MIT Press 1987). Recent pushback against the Putnam–Kripke outlook is found in ‘Natural Kinds and Natural Kind Terms: Myth and Reality’, Sören Häggqvist and Åsa Wikforss (*British Journal for the Philosophy of Science* 69/4: 911–933, 2018). For Two-Dimensionalism, see the page maintained by David Chalmers <https://consc.net/papers/twodim.html>, or his 2006b, ‘Two-Dimensional Semantics, in the *Oxford Handbook of Philosophy of Language* (E. Lepore and B. Smith, eds.; Oxford: Oxford University Press: 575–606, 2006.) For Frege on demonstratives and indexicals, see Robert May’s ‘Frege on Indexicals’ (*The Philosophical Review*, 115/4: 487–516, 2006).

● NOTES

- 1 Although it is debateable that a thing's necessarily being thus-and-so establishes that its being thus-and-so is part of the thing's essence. For example, it is a necessary property of each object that it be a member of its own unit class – of e.g. Madonna that she be a member of the class {Madonna}. But it is not part of her essence, the argument runs. See Fine (1994) 'Essence and Modality'.
- 2 Some natural kind terms are not indexicals. For example, elements such as *Berke-
lium* were defined before any examples had been found or synthesised.
- 3 We are exploring Kaplan's celebrated theory of indexicals, but with a simplifying change: we are pretending the theory is one of *utterances*, but actually Kaplan's theory assigns what he calls 'contents' – objects, extensions and propositions – to *expressions-relative-to-contexts*.

9

• pragmatics

Philosophers often talk as if the stating of facts or conveying of descriptive information were the sole purpose of language. Of course, it isn't. There are also questions, commands, predictions, greetings, jokes, story-tellings and many others, only some of which can be appropriately assessed as true or false. This territory was philosophically relatively uncharted before the past 60 years or so, when some philosophers – especially John Austin, H. P. Grice and John Searle – sought a theoretical framework in terms of which to describe it. Perhaps motivated by the later Wittgenstein's emphasis on the practical dimension of language – on its *use* for certain *purposes* – these figures have been so successful that the basic framework that emerged from their work has become a central part of mainstream philosophy of language and linguistics. In this chapter, we will explore its initial forms and motivations and consider some applications.

• MOOD AND FORCE REVISITED

In the Introduction, we distinguished *meaning* from *force*. Semantics concerns the former, and pragmatics is largely concerned with the latter. The sentences in Table 9.1, we said, *express the same proposition* but would normally be used to express it with different *force*. The grammatical differences are differences of *mood*.

It is natural to think that force is at least partly a psychological matter: in central cases, a person asks a question or makes an assertion only if they *intend* to. By contrast, the role of mood is clearly a matter of *convention*: for example, there is a convention whereby the use of the interrogative mood indicates to the listener that the speaker is asking a question – i.e. expressing the proposition with interrogative force. In English, the interrogative mood is typically accomplished by reversing the order of subject and verb (and to form the imperative in the simplest case, the subject of the command is made tacit, but is the intended audience).

Table 9.1 Mood and force

Sentence	Mood	Force (that would standardly be expressed by that mood)	Proposition (<i>content</i> , as it is often called in this context)
You are going to eat raw fish.	Indicative/ Declarative	Assertion	That you are going to eat raw fish.
Are you going to eat raw fish?	Interrogative	Question	That you are going to eat raw fish.
Eat raw fish!	Imperative	Command	That you are going to eat raw fish.

However, the relation of mood and force is somewhat fluid. The use of mood is not the only way to indicate force. The mere uttering of a sentence in a given mood is not sufficient for saying something with the force standardly conveyed by that mood. One can utter 'Eat raw fish!' without commanding, enjoining or suggesting anything. It is also not necessary. We can ask a question by using the indicative mood, usually by varying our tone of voice: using a rising intonation. That too is a convention; there is a convention according to which if you utter 'You are going to eat raw fish' with that sort of rising intonation, you are to be understood as asking a question, not as making an assertion. In written English, we can indicate a question by attaching a question mark to a sentence in the indicative mood, as in 'You are going to eat raw fish?'

● SPEECH-ACT THEORY

Yet another way is to say something like:

I ask you whether you are going to eat raw fish.

Indeed, instead of the rather complicated conventions in English by which we express force, we *could* eliminate distinctions of mood entirely from the language and employ a set of operators like 'I ask you whether...':

I ask you whether you are going to eat raw fish.

I assert that you are going to eat raw fish.

I command that you are going to eat raw fish.

Thus one might think that all moods collapse into the declarative. In his famous paper 'Performative Utterances' (1961), John Austin noticed that such forms do exist in English, but furthermore that their grammatical structure can be misleading,

at least to a philosopher. Consider the second one, 'I ask you whether you are going to eat raw fish'. Superficially, one might say that this is really an *assertion*, namely an assertion *that the speaker is asking the listener whether they are going to eat raw fish*. This, Austin held, is a confusion. The purpose of 'I ask you' is not to *describe the speaker*; when you utter 'I ask you whether you are going to eat raw fish', your aim is not to describe yourself. For if it were, it looks as if you would be describing yourself falsely, and a wag could say: 'No you're not!' (on the grounds that the sentence is declarative, not interrogative). The purpose of such an utterance is to *bring it about* that the speaker asks a question whose content is expressed by the words that come after. This is clear from the fact that someone making that utterance, in the right circumstance, does thereby succeed in asking a question.

Austin's idea is that there are certain sorts of words such that, in saying them, one thereby performs a certain kind of act – the act which, on a more superficial view, they might seem to describe or announce. The point is clearer from examples like this:

I apologise.

I hereby bequeath my cigar case to my nephew.

I promise to do that.

I name this ship the *Queen Elizabeth*.

Provided that the circumstances of utterance do not render the act 'void', as Austin called it (the words are not uttered in the performance of a play, the person is empowered to name the ship etc.), to utter such words is to perform the act of apologising, bequeathing, promising, naming. In committing the verbal act of uttering them, one does not *report the fact* that one performs the act – as if the act itself consisted in some other, perhaps nonverbal act. Rather, performing the act is actually constituted by the utterance. Austin, then, called these forms of words – I apologise, I bequeath, I promise, I name – *performative* verbs. The word 'hereby' is a good clue as to what is going on.

When he wrote 'Performatives', Austin was thinking that linguistic utterances can be divided into performative and non-performative (what he called 'constative'): a distinction between utterances that – according to some conventions – *do* something (perform an act), and those that say something, or more accurately *state* or *assert* something.

In practice, however, it proved difficult to describe this difference.

One problem is that what seems to be the distinctive feature of statements – their truth-evaluability – doesn't make anything like a clean cut between statements and performatives. On the one hand, performatives do 'confront facts', as in the possibility that one issues an *unjustified* warning or makes a promise which *one knows it is impossible* to carry out. On the other, for many statements, describing them as 'true' or 'false' is too crude; for example: 'Lord Raglan won the battle'; Raglan was the commanding officer but it was a 'soldier's battle', meaning that the outcome

had little to do with strategy etc. To simply respond to the statement with ‘True!’ or ‘False!’ would just be crude, papering over the distinctions that matter.

A second problem is that although *I assert*, *I ask* and so on do seem to be performative verbs, the saying of ‘I assert that my client is innocent’ and ‘My client is innocent’ seem equally to serve for an assertion that the client is innocent – for saying that the client is innocent. Likewise ‘Shut the door!’ and ‘I command you to shut the door’ serve to issue the command to shut the door. What, then, are we to say of the distinction between doing and saying? What this suggests is that saying should *not* be distinguished from doing; indeed, saying *is* a kind of doing, even if an especially important kind. This idea is reinforced by the observation that there are *many* kinds of acts that are acts of stating: saying, reminding, telling, informing, describing, criticising, alerting and so on.

Accordingly, Austin soon devised the theory he is most famous for, and which is now more or less the standard theory of what he called *speech-acts*. A normal speech-act comprises three main sub-actions (there is also a fourth, namely the *phonic* act of uttering certain phonemes; but we will pass over this). Sticking with our example of the question concerning the eating of raw fish, look at Table 9.2.

Table 9.2 Force distinctions within a speech-act

Locutionary act	The expression of content (meaning – cognitive and/or expressive)	Expressing the proposition that the intended audience is going to eat raw fish
Illocutionary act	The act performed <i>in</i> speaking (especially the expression of <i>force</i> , aka illocutionary force)	For example, asking whether the proposition is true
Perlocutionary act	The acts performed by the speaker <i>by means of</i> the illocutionary act that is an effect on the audience	For example, getting the listener to say whether or not they are going to eat raw fish

This is the central framework of speech-act theory as pioneered by Austin and developed by John Searle. Thus take the question, *Are you going to eat raw fish?* The speaker performs the **locutionary act** of expressing the proposition or content *that you are going to eat raw fish*. He expresses it by means of a sentence in the interrogative mood (Austin himself called this the act of “saying something” in the full normal sense’). Assuming the circumstances are appropriate – for example, the listener was in a position to hear – he thereby performs the **illocutionary act** of *asking the listener* whether they are going to eat raw fish. The **perlocutionary act** depends on the response of the listener; if the act is successful, it will be something like *getting the listener to say* whether or not they are going to eat raw fish.

Semantics is concerned with locutionary acts, with classifying sorts of expressions according to their meaning and with describing the meaning of various types of

expression. Locutionary acts *are not* confined to sentences that express propositions. There is a convention, for example, according to which ‘Ouch!’ expresses pain on the part of the speaker; but the meaning is non-propositional and wholly expressive rather than cognitive. The meaning expresses a state of the speaker, but does not describe anything. Similarly for the locutions ‘Blimey!’, ‘Rats!’ and the like. Yet these are not simply *noises* that are *caused* by pains and errors, as when a cat shrieks when you step on its tail. That ‘Ouch’ expresses pain is a convention of English, as you will quickly appreciate if you step on a foreigner’s toe.

Pragmatics is concerned in the first instance with illocutionary acts. According to Austin, illocutionary force is largely conventional: there exist certain rules or standards that in general one must meet in order to accomplish the relevant kind of illocutionary act. Cases of failed attempts at illocutionary acts are typically cases where a **presupposition** of the act fails to hold, as when you promise to pay a fine, when unbeknownst to you it has already been cancelled (we’ll say more about *presupposition* shortly). In such a case the circumstances for you to incur a commitment are lacking and the act is not performed. In the case of a successful accomplishment of an illocutionary act, the falsity of assertions is just one variety of Austinean ‘infelicity’. There are insincere assertions or promises, knowingly misleading warnings, and so on. Interestingly, in some cases, one cannot claim not to have performed the action on the grounds that one did not ‘mean it’ – i.e. did not have such-and-such intentions. Apologising by saying ‘I apologise’, for example, is a bit like bowing: one can do it while feeling nothing but resentment, but one does thereby apologise or bow all the same.

There are a great many illocutionary acts besides the basic ones just mentioned. One can *admonish, describe, warn, order, request, criticise, censure, welcome, reprove, compliment, greet, berate, plead, chastise* and so on (some types will include others as sub-varieties). John Searle proposes that the range of basic illocutionary acts be set at five, namely the *assertive, directive, commissive, declaratory* and the *expressive*, characterised as follows: the *assertive* is to say how things stand in the world (as is often usefully put, the ‘direction of fit’ is word to world); the *directive* (which includes the interrogative) is trying to get hearers to do something (the direction of fit is world to word); the *commissive* – ‘I will ...’ etc. – to commit themselves to action (world to word); the *declaratory* – like Austin’s performative utterances – is when they perform some action solely by virtue of saying so (both directions of fit; declaratives ‘create statuses’, and make possible institutional facts: legal obligations, property, marriage, money etc.); the *expressive* when they intentionally make manifest their attitudes (word to world; see Chapter 14, ‘Slurs’).

The performance of a locutionary act is not sufficient for the performance of an illocutionary act. A person without the necessary authority cannot marry a couple by saying ‘I now pronounce you married’. If the person cannot hear you, you can try but fail to warn the person. Arguably, there are ‘uptake’ conditions even on the speech-act of assertion: you attempt to tell a person something, and you did utter the intended words, but the person cannot understand what you are saying; your act of telling is not successful. Or the fault may be more insidious: it may be accepted amongst certain

people that the right of a person of a certain age, sex or class – a teenager, a female, a servant – to refuse something, to claim something, or even to speak, may be denied. For example, no amount of saying ‘No!’ would count in their eyes as a genuine refusal; if uptake is indeed required, then no speech-act of the intended kind can be performed. This is the phenomenon of **illocutionary silencing**, which can take many forms, from the subtle to the not so subtle.

The study of *perlocutionary* acts would not be an enquiry into the conventions of language so much as the psychology and social features of actual linguistic episodes. The perlocutionary act performed by means of an illocutionary act involves an effect on the audience: it must be describable in terms of a transitive verb, where the subject of the verb is the speaker and the object is the audience. Thus, for example, ‘In speaking that way, General Patton roused the troops’ would be a description of a perlocutionary act; one can also *anger* one’s audience, *persuade*, *interest*, *frighten*, *bore* them and many others. Again, observe the effect of ‘hereby’: unlike ‘I hereby ask you to shut the gate’, ‘I hereby persuade you to shut the gate’ is a misuse of ‘hereby’. Perlocutionary acts do not have to be *intentional*, but they do have to be describable as things that the speaker *did*. Thus, for example, Patton may have intended by his speech to rouse the troops, but succeeded only in *frightening* them. Though not intentional, *frightening the troops* would still be something that General Patton *did*, and did by speaking as he did.

There is no end to the variety of possible perlocutionary acts, and they are not in general governed by convention in the way that illocutionary acts are. But they are essential to an understanding of language nevertheless because speaking is an *intentional activity*. We don’t just speak for the hell of it; we speak to people for reasons, namely in order to have some effect on the people we speak to. A perlocutionary act is typically the end, or goal, we are shooting for in speaking as we do, even if we fail and end up performing some other perlocutionary act or none at all.

● IMPLICATURE

It is natural to think that *conveying information* is entirely a matter of semantics. But in fact, information is often not only conveyed, but intentionally conveyed by utterances which are *not* explicitly encoded by the utterance involved, which is not strictly speaking part of the semantics of the utterance.

Consider the famous example due to Paul Grice. Suppose that Jones’ parents ask Headmaster Smith how Master Jones is getting on at the boarding school. Smith replies, ‘Well, he hasn’t been expelled’. Mr. and Mrs. Jones would naturally take Smith to intend the message that Master Jones is not doing well at school, not well at all. Headmaster Smith did not *say* this, and what he says does not *logically entail* it either, for it is not inconceivable that what Smith said should be true yet Master Jones is doing well at school. But, nevertheless, Headmaster Smith did succeed in communicating the message that he is not doing well; Mr. and Mrs. Jones would be right in concluding that their son is not doing well. Grice introduced the term

conversational implicature for this phenomenon, or ‘implicature’ for short: Smith’s statement *implicated* that Master Jones is not doing well at school.

In general:

A statement P *implicates* the proposition Q if: P does not logically entail Q, but a well-informed witness to the utterance, who is a competent speaker of the relevant language, would take the speaker to be intending to convey Q.

To mix Austin’s terminology with Grice’s, implicature is a kind of intended perlocutionary force.

One important complication: as above, it is at least possible that what Smith said should be true when Master Jones *is* doing well at school, yet the Headmaster did not intend the implicature – when indeed he spoke rightly. For conversational implicature always admits of being *cancelled*: for example, if the Headmaster adds ‘... just in case you were worried – as you have no doubt heard, nearly half his classmates were confined to their rooms yesterday due to the unpleasant events that took place in the refectory’.¹ Only if the Headmaster does not cancel the implicature is it reasonable to understand him in the way indicated.

Suppose instead that Headmaster Smith were to have said: ‘Well, he hasn’t been expelled’ when he knew that Master Jones was doing well, yet he did not say anything to cancel the implicature. He would not have said anything false. It would not be *semantically* improper for him to have said this. What he said, however, would have been *pragmatically* improper. In fact, it is precisely because it would be pragmatically improper for him to say this, under such circumstances, that it would be reasonable for the listener to understand him to be conveying that Master Jones is not doing well.

The reason, according to Grice, is that there exists a certain general maxim of conversation that he calls the **cooperative principle**. This principle comprises several sub-maxims:

- **Maxims of Quality:** do not say what you believe to be false; do not say anything for which you lack adequate evidence.
- **Maxims of Manner:** be brief; avoid obscurity, ambiguity, vagueness.
- **Maxim of Relation:** be relevant.
- **Maxims of Quantity:** make what you say neither more nor less informative than is required for the purposes at hand.

Conversation is, in this sense, a *practice*, that is, an activity implicitly involving certain rules. Adherence to these rules, however, is not all-or-nothing; they can be obeyed or flouted in various ways and to varying degrees. Conversation is not a tightly defined activity, and different forms of conversation – banter, theoretical

discussions, debates, arguments, heart-to-heart talks, irony, chat and so on – each have their own weightings of these rules, perhaps along with rules or principles of their own.

Still, the bottom line is that people in conversation of various kinds understand each other to be cooperative: as trying to help one another. Thus the reason that conversational implicature works is that speakers understand each other to be trying to obey the sub-maxims of the cooperative principle (in very much the same way that game players understand each other as trying to win the game). Thus Jones' parents would normally understand Smith to be trying to be informative. Thus, if Master Jones were doing well, Smith would have said so: not being expelled is a necessary condition of doing well, but it is far from sufficient, and if Master Jones were doing well, there would be far more to say about it than that. So it would be natural and justified for Mr and Mrs Jones to reason: Smith is trying to be informative and truthful, but he is also trying to be polite or avoid hurting our feelings; so he is saying the most positive thing he can say; since that is the most positive thing he can say, he must be trying to convey that our boy is not doing well.

● SOME APPLICATIONS OF THE CONCEPT OF IMPLICATURE

The notion of implicature can help to explain a number of phenomena that would otherwise obstruct or complicate the logical analysis of language. For example, according to the standard truth-functional analysis of the conditional 'if-then' of English, the following statement would be true:

If donkeys are smaller than fleas, then donkeys are larger than cats.

The reason is that according to the truth-functional analysis, a conditional with a false antecedent is *true* (see Chapter 1, 'Logical syntax and logical operators'). This is bizarre, since the conditional seems quite wrong, since obviously *nothing* is smaller than a flea but larger than a cat.

The impropriety of such a conditional, however, can now be explained in pragmatic rather than semantic terms. It would indeed be wrong to *say* it. But what makes it wrong to say it is not its failure to be true.

Of course, everybody knows the obvious fact that donkeys are not smaller than fleas. Therefore the conditional is *not informative* to anyone. One knows that donkeys are horse-like animals, much bigger than fleas; one therefore knows that a case when the antecedent is true, and the consequent is also true, will never happen, just because the antecedent is always false. Since, according to the truth-functional analysis, the falsity of the antecedent entails the truth of the conditional (but not the other way around), an assertion of this conditional violates the maxim of quality, which is the rule of informativeness: a person cannot learn anything from it. Thus the conditional is true but **pragmatically improper**.

In general, an assertion of a conditional statement is appropriate (i.e. informative and relevant) only if the truth-value of both antecedent and consequent are *unknown* (or not explicitly known) to at least some of the participants in the linguistic transaction, or can reasonably be presumed so by the speaker. Thus in typical circumstances a conditional whose antecedent's truth-value is explicitly known to be false is *not normally assertible*. So no wonder the donkey-conditional looks so weird.²

But if it is not the truth-values of the constituents that justify the statement, what does? Answer: some other kind of *connection* between them, often but not always a causal connection. For example, one might say 'If the pork was off, then the dog will throw up' – not being certain whether the pork was off or whether the dog is going to throw up, but believing that dogs usually vomit bad meat. If one accepts the conditional, and if one subsequently finds the antecedent to be true, then one can infer the consequent (or if one accepts the conditional, and finds the consequent false, one can infer the falsity of the antecedent). There are many exceptions, but in general an assertion of a conditional is proper only if the speaker reasonably assumes the truth-values of its antecedent and consequent are unknown or not explicitly known, but has some other grounds for asserting the conditional. Otherwise the cooperative principle is violated.

Nevertheless, we do have jokes like 'If that's a jackdaw then I'm a monkey's uncle', said by someone who knows that it's not a jackdaw. Since everyone knows that he isn't a monkey's uncle, everyone understands him to be saying that it's not a jackdaw. The possibility of such jokes suggests that we do implicitly understand the conditional truth-functionally.

Violation of the cooperative principle – in particular the maxim of quantity – explains why, if one knows, for example, that Marie's car is a Fiat, it would be wrong to say 'Marie's car is a Fiat or it's a Volkswagen', despite the fact that the disjunction is true. It would be misleading, suggesting that one holds the two options open.³ Likewise to say 'All mammals bigger than blue whales are nocturnal' is, according to the classical account of quantification, true, because there are no mammals bigger than blue whales; but to say so would be misleading, since one could assert the more informative 'No mammal is bigger than a blue whale'. And there are many other examples where the Gricean strategy promises to remove the apparent tension between intuition and orthodox semantics or logic.

It is largely because we do implicitly rely on cooperativeness that we are able to convey so much information with so few words. For example, given the information jointly available in the context, one could convey that one won't come to this evening's party by saying: 'I'm in Edinburgh' (the party is London – it's too much hassle to come just for a party etc.). Or one could decline to go to a party by saying 'I'm studying for an exam in the morning'.

However, this points to a complication with the Gricean point of view. So far, one might think that the conversational maxims come into play only once one knows what propositions are expressed by an utterance. The maxims are triggered by the need to discover the implicatures of an utterance, not by the need to identify the proposition literally expressed by the utterance. But it seems that semantics and

pragmatics do often necessarily interpenetrate; in such cases the two must work together, as a team. For example, suppose your roommate says:

The landlord said on Tuesday he would fix the window.

The sentence is ambiguous: is the proposition that the landlord's statement that he'd fix the window happened on Tuesday, or is it that the landlord said that Tuesday is the day that he'd fix it? Typically, one would rely on information about the context, including information relevant to the speaker's likely beliefs and purposes to disambiguate such an utterance, to find out which proposition it expressed. You might be in a position to reason: 'my roommate and I both know that the landlord works full-time, and never comes to us except on weekends; so the first interpretation is more likely'.

So perhaps the identification of the propositions expressed by an utterance and of the implicatures of the utterance is better thought of as different aspects of a single process of interpretation: one which takes stock of all interpretation-relevant information.

● PRESUPPOSITION: STRAWSON'S AND DONNELLAN'S OBJECTIONS TO RUSSELL'S THEORY OF DESCRIPTIONS

Strawson. According to Russell's theory of definite descriptions (Chapter 3), definite descriptions – expressions of the form 'the F' – do not have the semantical function of referring, because their significance does not depend on the existence of a unique F or of Fs generally. Instead, a sentence such as 'The King of America is rich' says that there is a unique person who is the King of America, who is rich. The proposition expressed by the sentence is therefore false, since there is no such thing as the King of America.

P. F. Strawson (1919–2006), in his famous 'On Referring' (1950), voiced the following objection. If someone were actually to say 'The King of America is rich!', we would not react by saying that what he said is *false*. We would not say 'No, the king of America is not rich!' We would say something like, 'No, no, you're under a misapprehension; America has no king.' The whole business of saying something truth-evaluable has not got started. Such a person, according to Strawson, has attempted to *use* the phrase 'the King of America' to *refer*, but failed, as if they were aiming blindfolded to strike a piñata when there is no piñata. According to Strawson, the use of definite descriptions **presupposes** the existence of their referents. To use them does not normally amount to *saying* that their referents exist.

The cash value of the distinction between presupposition and entailment is this:

If P entails Q, then: if Q is false, then P is false.

If P presupposes Q, then: if Q is false, then P is neither true nor false.

Crucially, Strawson holds that the meaning of a *referring expression* of the form 'the F' does not depend on whether it is used successfully to refer. Whereas meaning, for Strawson, is a feature of expressions, reference is a feature of speech-acts, of *uses* of expressions. Referring is something we *do*, not a property of expressions considered in the abstract. In general, the reference of any singular term (including definite descriptions) is presupposed, *not* entailed, by their standard use in sentences. To conceive reference as a feature of utterances rather than of linguistic expressions is indeed intuitive, counting Noam Chomsky amongst its adherents.

In effect, Strawson argues for a return to Fregean ideas: the meaning of 'The F is G' is such that if there is no such thing as the F, the meaning of the sentence is unaffected, but, nevertheless, it cannot be used to say anything with a truth-value. Meaning for Strawson looks a lot like Fregean sense. Therefore, at least at first blush, he will fall into the problems discussed in Chapter 2 surrounding true claims of nonexistence, such as 'The King of America does not exist'. However, his emphasis on the concrete use of language rather than the rarefied air of functions, propositions and senses brings him closer to the practice of ordinary language; it's not for nothing that together with Austin he is routinely put forth as a quintessential *ordinary language philosopher*. They've made a plausible case that by sticking closer to the surface phenomena of ordinary language, such problems as the puzzle over negative existentials are not intractable. And the idea of presupposition, so central to the approach, has been enormously influential in both philosophical and linguistic pragmatics.

Presupposition is in fact quite rampant, going well beyond presuppositions of reference. We've touched on the phenomenon before in connection with Austin. 'She regrets having kissed him' presupposes that she kissed him; 'He no longer plays the piano' presupposes that he used to play the piano; 'She stopped fiddling her taxes' presupposes that she once was in the habit of fiddling her taxes.

Donnellan. Another observation of the ordinary use of definite descriptions that seems to upset Russell's picture, but also upsets Strawson's, and which nevertheless does imply that at least sometimes they function as referring expressions, was made by Keith Donnellan. We are in the audience to see Jones being tried for Smith's murder. The real murderer, however, is Brown, who is a very sane man, a cold-blooded killer as we say. But no one, save Brown, knows that Brown killed Smith, and in fact you and I are both persuaded that Jones killed Smith. This is partly because of Jones's egregiously bizarre behaviour on the witness stand. In fact, Jones is insane. I whisper to you, 'Smith's murderer is insane!' to which you reply, 'Yes, I think so'.

It seems plain that I did communicate to you a truth *about Jones*, in some way equivalent to 'Jones is insane'. Yet I did this by means of a sentence that, according to Russell, is false because it expresses a proposition that is about Brown, a sane man, not Jones – in the sense that Brown satisfies the description 'Smith's murderer' ('the person who murdered Smith'). Donnellan (1966) therefore distinguishes the **referential use** from the **attributive use** of definite descriptions: one's use of a definite description is *referential* if one uses it to enable the audience to identify the referent; one's use of a definite description is *attributive* if one says something about whoever or whatever satisfies it, just as Russell's view would have it. The key is that

a referential use involves the description only inessentially; it is merely a tool for communication, not part of what one communicates. It doesn't matter how one packages the message, so to speak, so long as one successfully communicates it. The trouble with Russell's view is that there seems to be no way to characterise such cases – such cases of a successful use of a description to refer to a thing when the thing does not satisfy the description.

The position outlined in Strawson's 1950 paper cannot accommodate the distinction either: Strawson's theory agrees with Russell's in that if the F exists, then a use of a term of the form 'the F' simply refers to the F.

In the above example, I spoke only vaguely of 'what he communicated' rather than the 'proposition expressed by' my utterance of 'Smith's murderer is insane'. Does that sentence, in that context, express a false proposition about Brown or true proposition about Jones? If we say the former, then perhaps there is no objection to Russell here, but the account will tell us nothing about the actual linguistic transaction. But if the latter, then it seems that you, once apprised of the facts of the case, would have to agree that my utterance of 'Smith's murderer is insane' is somehow true – equivalent to 'Jones is insane' – even though you know that Jones is *not* Smith's murderer (Brown is).

Kripke. Kripke urged in 1977 that the matter can be cleared up, leaving Russell's theory relatively unscathed, by invoking an important distinction closely related to Gricean distinctions. If, during a heist, one burglar says to his partner, 'The cops are around the corner', then it is evident that what is intended is different from the literal meaning of the sentence used in the particular context: he means something like 'We can't wait around collecting loot; let's go!' Thus *speaker's meaning* can be distinguished sharply from *semantic meaning*: whereas the latter is determined by linguistic rules governing the use of expressions, the former is determined by various special intentions of the speaker, including Grice's cooperative maxims (thus conversational *implicature* is most plausibly reckoned a matter of speaker meaning).

Similarly, Kripke distinguishes between *semantic* reference and *speaker's* reference. You and I see Bob in the distance, but we both mistake him for Jarda. I ask you, 'What is Jarda doing?'; 'Raking leaves', you reply. The semantic referent is Jarda, but the speaker's referent is Bob. It seems that we can say exactly analogous things about the case of Smith's murderer: the semantic referent is Brown, the speaker's referent is Jones. The difference, then, is a pragmatic difference, not a semantical one – that is, it is not something that would affect the analysis of referring terms or definite descriptions. Donnellan's phenomenon is no more mysterious than the fact that tools can unintentionally but successfully be misused, as when a screwdriver is used as a chisel by someone who thinks that that is what it's for.

● METAPHOR

The subject of metaphor is a relatively untamed beast, which is liable to lead us into a dizzying variety of issues in psychology, epistemology, literary criticism, aesthetics,

philosophy of science, philosophy of mind and other things. Here, I just say enough to be provocative and to convince you that the tools we have been introduced to shed some light on the phenomenon; beyond that I refer you to the further reading.

Here are some metaphors:

Richard is a lion.

Life is a yo-yo.

What he said casts some light on the matter.

She was carried away by passion.

Light consists of waves.

He burns for her.

I've some new ideas floating around in my head.

She was the finest flower of her generation.

I'm on a highway to hell.

Money doesn't grow on trees.

Giovanni followed in his father's footsteps.

What is a metaphor? Many initially plausible but on reflection not-so-plausible things have been said about this. For example:

- A metaphor states a *comparison*. No; it doesn't *state* a comparison; 'New York is bigger than Paris' states a comparison. At most, the metaphor implies or suggests a comparison.
- Metaphors are *similes*. No; metaphors typically are more powerful than their corresponding similes; 'Richard is a lion' versus 'Richard is like a lion'. Unlike the metaphor, the corresponding similes are typically true.
- A metaphor states a *similarity*. No; *similes* do that, such as 'My love is like a rose', as opposed to 'My love is a rose'.
- Metaphors are literally false. No; 'Money doesn't grow on trees' is literally true.
- A metaphor is equivalent to a conjunction of literal sentences. No; no conjunction of sentences is equivalent to Seamus Heaney's 'potatoes piled in pits, blind-eyed'.
- In a metaphor, the meanings of words shift from literal meanings to metaphorical meanings. No; 'Richard is a lion' *could not* have the metaphorical significance that it does if 'lion' were not understood as denoting the usual African felines.

Metaphor is easier to think about if we view the property of *being metaphorical* as a matter of pragmatics. This means that it is not *sentences* or *propositions* that are metaphorical; rather it is certain *speech-acts* – even though the sentences will typically be without a reasonable literal use. We should try to characterise metaphor as a kind of use of language. Metaphor is not a locutionary matter and, we'll see in a moment, there are even doubts about whether it is an illocutionary matter.

Here are some elements that should be incorporated into any account of metaphor; the picture that will emerge is more or less a Gricean one.

- (1) A correct metaphorical utterance would violate Grice's cooperative principle if taken simply and literally (that is, as expressing a statement whose content is the conventional meaning of the sentence). Thus, for example, 'She burns for him' would hardly make sense; 'Richard is a lion' would too obviously be false; 'Money does not grow on trees' would too obviously be true. Thus the listener, expecting the speaker to obey the cooperative principle – expecting the speaker to have a reasonable purpose in speaking – is prompted to take the utterance in some other way.
- (2) Nevertheless, metaphorical significance depends on literal meaning: as before, 'Richard is a lion' *could not* have the metaphorical significance that it does if 'lion' were not understood as denoting the usual African felines.
- (3) Metaphors may have a *cognitive* function. For example, it is scientifically useful and informative to think of light as a wave, even if literally it cannot be a wave; a wave is a moving region of compression in a physical medium (water, air etc.), and light can propagate through a vacuum.
- (4) Metaphors may also have an *expressive* function (a function of language that does not affect the truth-conditions of what is said). In poetry, or when we use metaphors to describe how we feel, think or perceive, we often do so in order to convey *what it is like* to feel, think or perceive as we do.

(3) and (4) are crucial. We have seen already that literal language has both cognitive (representational) and expressive functions; so it is with metaphor.

It is plausible that what these functions have in common is the involvement of the imagination. For a cognitive example, we imagine tiny objects moving around a larger object at the centre; such proved a useful model for the *atom*, even though the atom is not actually like that. For a non-cognitive, expressive example, we imagine the pain of being stabbed; we thus get the point of someone's saying 'I realised with a stab that she was being unfaithful to me'. For an example that is perhaps both, we imagine a beautiful girl, and also the sun; as I. A. Richards puts it, the distinctive 'feel' of Romeo's 'Juliet is the sun' arises from an 'interaction' among those acts of the imagination. It may be impossible to sum up all such activities of the imagination. What we can say, however, is that it is the overarching purpose of the metaphorical utterance to stimulate such activities in the imagination of the audience. In the terminology of speech-act theory, metaphor is characterised by a type of *intended perlocutionary act*.

● HISTORICAL NOTES

Ludwig Wittgenstein's later work – both his famous *Philosophical Investigations* of 1953, and *The Blue and Brown Books* of the 1930s, which was unpublished during his lifetime but circulated unofficially in mimeographed form around Cambridge in the 1930s – played a decisive role in the rise of so-called 'ordinary language'

philosophy, which reached its peak in the 1950s and 1960s. Gilbert Ryle (1900–76), Peter Strawson (1919–2006), John Austin (1911–60), Norman Malcolm (1911–90), O. K. Bouwsma (1898–1978) and many others applied the subtleties of careful observation of ordinary language to philosophical problems with sometimes striking results. The legacy lives on in that philosophers as a whole are usually much more sensitive to how philosophical theories are so seldom reflected unambiguously in ordinary linguistic practice (in ‘what we say’) than previously. But undoubtedly a much greater effect of ordinary language philosophy was its invigoration of pragmatics – understood as a sub-discipline of linguistic study alongside syntax and semantics, though sometimes contending with them. By 1970, Austin’s work in particular had inspired Searle in writing his *Speech Acts: An Essay in the Philosophy of Language* (1969), and then his *Expression and Meaning: Studies in the Theory of Speech Acts* (1979); meanwhile the influence of Grice (1913–88), especially his essay ‘Logic and Conversation’ (1975), began to make itself felt. Indeed, although we haven’t considered it here, Grice can be viewed as proposing an alternative, full-fledged model of speech-acts to Austin and Searle’s, replacing their use of *convention* with the speaker’s and hearer’s intention (‘Meaning’ [1957], in Grice 1989). The Austin–Searle and Gricean paradigms have since then been subjected to massive development, revision and also pressure: (1) Stephen Schiffer and others have developed Grice’s more general ideas – his account of what are essentially speech-acts – into ‘Intention-Based Semantics’, which fits with certain theories in cognitive science and the philosophy of mind (see his *Meaning*, Oxford University Press, 1972); (2) the literature on implicature in particular has grown enormously in the past 30 years, with various types of implicature being identified, new maxims being formulated and new ways to cope with cases in which the original maxims are found wanting; (3) according to relevance theory, proposed by Dan Sperber and Deirdre Wilson, Grice’s maxims should be replaced by a single, general principle of optimal relevance or communicative efficiency: roughly, given at the outset a speaker’s understanding of the context, including information about the likely cognitive states of the listeners, the speaker says only enough to affect the audience’s cognitive states in the desired way. Grice’s theory and relevance theory conflict in certain cases, and not always in favour of relevance theory; but relevance theory is an intuitive, flexible and compact alternative to Gricean maxims. The phenomenon of illocutionary silencing gained attention with Catherine Mackinnon’s book *Feminism unmodified. Discourses on life and law* (1987) and received attention from the philosophy of language with the work of Jennifer Hornsby (1993) and Rae Langton (1993). Closely related is the phenomenon of hermeneutical injustice, as analysed in Miranda Fricker’s (2007). See Chapter 14, ‘Conceptual Engineering’.

● CHAPTER SUMMARY

The mood of a sentence is a syntactical feature that is conventionally tied to the type of force that may be attached to an utterance of the sentence. For example, switching the order of subject and verb – e.g. ‘you are’ versus ‘are you’ – is a standard difference between the declarative and interrogative moods. Speakers sometimes break the principle, as when asking a question with the declarative mood.

The standard force-types – assertions, questions, commands – give rise to a more fine-grained theory. If what we are considering is a speech-act of assertion, Austin famously distinguishes the locutionary act of expressing a proposition from the illocutionary act of assertion of a proposition, and from the perlocutionary act of getting the listener to believe a proposition. Locutionary and illocutionary acts depend mostly on the actions of the speaker; the perlocutionary act depends entirely on the effect on the listener's mind. The precise extent to which illocutionary acts are conventional as opposed to intentional is a matter of dispute. Normally, we intend the perlocutionary act we succeed in performing, but often the actual perlocutionary act is not what we intend. The perlocutionary act, however, is always something that the speaker did in speaking as they did, whether or not it was intentional.

A statement in which P *implicates* Q is: P does not logically entail Q, but a well-informed, competent speaker would take the speaker to be intending to convey that Q. This captures the ordinary sense in which we sometimes say: 'I know you didn't actually *say* that he's ugly, but you *implied* it'. There are no strict rules for what is known as conversational implicature, but the idea is that the practice of conversation possesses the following constitutive rules, known collectively as the cooperative principle: do not say what you believe to be false; do not say anything for which you lack adequate evidence; avoid obscurity, ambiguity, vagueness; be brief; be relevant; make what you say neither more nor less informative than is required for the purposes at hand. Working out conversational implicatures involves the assumption of fidelity on the part of the speaker to the rules of conversation; a speaker may not have said anything that logically entails Q, but the assumption that the speaker aims to satisfy the cooperative principle may imply that they did mean to convey Q.

Another key concept is that of *presupposition*: if P presupposes Q, then if Q is false, P is neither true nor false (whereas if P entails Q, then if Q is false, P is false). This formed the basis of Strawson's attack on the theory of definite descriptions: if there is no unique F, rather than saying as Russell does that 'The F is G' is false, we can say that 'The F is G' presupposes reference on the part of the speaker's use of 'the F', and that since that presupposition is false, the use of the sentence does not express a statement with a truth-value. Donnellan presses harder, pointing out that 'The F is G' can serve to express a true statement, even when it is false taken in Russell's sense; he calls the two readings the referential sense and the attributive sense. Kripke in effect saves Russell, suggesting that *speaker's reference* can differ from *semantic reference*; both Strawson's and Donnellan's phenomena can be explained purely pragmatically, with no threat to Russell.

There are many theories of metaphor, but it is plausible to regard it as strictly a pragmatic phenomenon – in particular that it will violate maxims of cooperativeness if taken literally (according to semantic rules) and it always involves the dominant intention to achieve a perlocutionary effect of stimulating the audience's imagination in certain distinctive ways. Thus the semantic or literal meaning may be false and it may be true, but the point is always to engage the audience's imagination. The difficult work is to describe these effects on the imagination.

● STUDY QUESTIONS

- 1 Eavesdrop on a conversation – unobtrusively! – and try to identify the various illocutionary acts that you hear. How many of them sit comfortably in one of the basic categories of *force* we identified in ‘Mood and force revisited’, at the start of this chapter? And describe the perlocutionary acts you hear.
- 2 What, if anything, would be wrong with following claim: ‘The purpose of an utterance is always the same: to express one’s mental state!’?
- 3 Is it really true that every type of (illocutionary) speech-act is conventional? Compare asserting, promising and saying ‘I do’ when getting married.
- 4 To work out by Gricean means conversational implicatures seems rather complicated; what would a defender of Grice say to the charge that the complicated and recondite procedures of linguistic analysis he describes cannot possibly take place in ordinary conversation?
- 5 Consider:
I ate some of the cake.
Suppose it turns out that the speaker ate *all* of the cake; has the speaker lied (has the speaker wilfully uttered a sentence which they know is false)?
- 6 ‘Money doesn’t grow on trees’ – literally true, therefore not all metaphors are literally false. But are there literally true metaphors that do *not* contain negation? What of ‘Business is business’? Is that a metaphor? How does an utterance of it achieve its intended effect?

● PRIMARY READING

- Austin, J. L. (1962) *How To Do Things With Words*, second edition.
 Donnellan, K. S. (1966) ‘Reference and Definite Descriptions’.
 Grice, H. P. (1989) ‘Logic and Conversation’ [1975], in his *Studies in the Way of Words*, pp. 22–40.
 Kripke, S. (1977) ‘Speaker’s Reference and Semantic Reference’.
 Searle, J. (1969) *Speech Acts: An Essay in the Philosophy of Language*.
 Searle, J. (1979) ‘Metaphor’.
 Strawson, P. F. (1950) ‘On Referring’.

● SECONDARY READING

- Harris, D. Fogal, D. and Moss, M. editors (2018) *New Work on Speech Acts*.
 Hills, D. ‘Metaphor’, in the *Stanford Encyclopedia of Philosophy*.
 Lepore, E. and Stone, M. (2014) *Imagination and Convention*.
 Sperber, D. and Wilson, D. (1995) *Relevance: Communication and Cognition*, Second Edition.

● NOTES

- 1 Grice distinguishes between *conversational* implicature and ***conventional implicature***. In the conventional case, the implicature is not *cancellable*: there are no

circumstances in which one could assert the sentence without being responsible for the implicature. An example is 'He hurried home', which always implicates that he had a reason to get home quickly, even though it would not be a logical contradiction to say 'He hurried home, but for no reason'. It is controversial whether conventional implicature exists, however. Especially when the same phenomenon might be covered by the notion of analytic implicature, which is just the notion of the analyticity of certain conditionals *If p then q*, i.e. *If he hurried home then he had reason to get home quickly*.

- 2 The case is different for subjunctive or counterfactual conditionals, like 'If dogs could talk, they would be honest' – in those cases we know the falsity of the antecedent. Unlike the material conditional if–then, the counterfactual conditional is not truth-functional: its truth-value is not determined by the truth-values of its antecedent and consequent.
- 3 The disjunction is truth-functionally equivalent to 'if Marie's car is not a Fiat then it's a Volkswagen'; since Marie's car is a Fiat, the conditional is true but pragmatically incorrect to assert. Likewise for 'if Marie's car is not a Volkswagen then it's a Fiat'.

10

• davidson's philosophy of language

With certain exceptions, the theories discussed so far do not attempt to define, in any sense, the concepts of meaning and reference. They *use* those concepts, albeit in ways that are much more disciplined than the ways they are used ordinarily, in our 'folk theory', our common-sense ways of talking about meaning. The finished concepts are sharper and more lucid than the corresponding folk-concepts but are in some sense implicit in them. Similarly, Newton's theory of gravity doesn't define the concept of gravity but contents itself with using a sharper version of it, relating the concept to others such as mass and distance. But as familiar as they are, the concepts or phenomena of reference and meaning can seem, at least to some, a little spooky. What sorts of phenomena are they? Are they like digestion or electrical charge? Or are they like something mathematical or some other kind of non-physical property? When I refer in conversation to, say, the highest placed pebble on Mt Kilimanjaro, by using those words, are there some sort of meaning-rays, or referential-rays, that connect my mind or the words to the object?

Donald Davidson endeavours to describe the facts underlying meaning and reference *without* simply using those concepts. He was largely inspired by Quine, the subject of Chapter 6. Quine propounds a form of naturalism, and is a noted sceptic about the concept of meaning employed by Frege, Russell and others; he proposes an alternative conception of what having language amounts to that does not simply assume the concept of meaning. Davidson agrees with Quine that linguistic phenomena should be investigated without making the sorts of presupposition found in Frege and Russell. And he agrees that *if* it is possible to theorise fruitfully about meaning, then the concept must be explained as empirically answerable, applicable to language-users from a third-person point of view. But he is more optimistic than Quine. He is, at any rate, much less inclined to say that if the theory of meaning cannot take its place among the hard sciences, then so much the worse for it. He thinks, roughly, that it is undeniable that meaning is something we *know* about, and thus that it must be possible to give a systematic description of what we know. And

he is encouraged in this thought by his rather remarkable success in showing how a theory of meaning *can* be made into something empirically answerable or inter-subjectively accessible.

● METHODOLOGY

Davidson's key concept is the idea of a *theory of meaning for L*, which is a theory that states all the semantical facts about a given natural language *L*, such as Portuguese. Ideally, for each meaningful expression of *L*, such as a sentence, the theory states its meaning, in the sense of *entailing a theorem* that states its meaning. The language that the theory of meaning is *about* is called the object language, and the language that we *use* to state the theory is called the metalanguage. If (i) we can say what a theory of meaning for *L* would be like for *variable L* – any language *L* you like to take – we thereby know the shape or form of a theory of meaning, and (ii) if we can say how, in general, such theories are to be confirmed 'in the field', then we have said what there is to say philosophically about meaning-in-general. General philosophical knowledge about meaning can thus be summed up by describing (i) the form of a theory of meaning and (ii) how such a theory is confirmed.

Unlike other sorts of phenomenon such as gravity or radioactivity, meaning is essentially something *known*, in the following sense: the meaning of every meaningful linguistic expression is (or has been) known to someone or is logically entailed by something that is or has been known to someone. But a theory of meaning, for a particular language such as a linguistic theorist might devise it, will likely be a rather technical beast, using linguistic concepts and so on of which the typical speaker of a language has no inkling. So Davidson envisages the relation between theories of meaning and common-sense knowledge of actual speakers in the following way: ordinary speakers cannot normally *explicitly* state a theory of meaning for the language they speak, but it is plausible to think of a theory of meaning as making explicit what the ordinary speaker *implicitly* knows. A theory of meaning describes what it is that an ordinary speaker implicitly knows.

● THE GENERAL FORM OF A THEORY OF MEANING

COMPOSITIONALITY

What would such a theory look like? First, such a theory must be *compositional* in the sense described in the Introduction of this book: it must show how the meaning of a sentence is determined by the meanings of its parts. Consider a speaker who knows the language *L*. Ignoring sentences that are too long or complicated, such a speaker understands each sentence of *L*, in the sense that if presented with it, they will know what it means. Beyond a certain point, sentences will be too long for

the speaker to understand; so the set of sentences that the speaker would actually understand is finite. Nevertheless, it seems undeniable that the finitude of this set is due only, so to speak, to limitations in the hardware and not to limitations in the programme. Where a sentence is simply too long for a given speaker to understand, the problem is not that the speaker does not (implicitly) know the principles that determine the meaning of the sentence, but that they are unable to apply them correctly.

To refresh you with a simple example, we all grasp the principle that determines the meaning of 'the father of X' for any singular term that we understand put for X. But X may itself be of the form 'the father of ...'; so we get the singular terms 'the father of the father of X', and 'the father of the father of the father of X' and so on. This reflexive behaviour – of the device accepting as inputs its own outputs – is what is known as **recursion**. From this device we can get sentences of arbitrary length – for example, of the form 'The father of the father of X was tall'. Beyond a certain point, we will be unable to interpret them, because we will become confused, sleepy or dead. Still, we do grasp principles that would enable us to understand them, if we were not contingently limited in the ways that we are.

The demand for compositionality also emerges in another way. Knowing what a sentence means, in the sense in which speakers know the meanings of sentences of their own language, is more than simply knowing *that s means that p*, for some *s* and *p*. For example, one might be reliably informed that a certain sequence of phonemes of Lao means that snow is white; it does not seem that you thereby *understand* the sentence of Lao. In order to understand it, one has not only to know what its meaning happens to be but also what its constituent words mean and how its meaning what it does depends on the meaning of those words and the way it is constructed from them. Understanding, hence knowledge of meaning in the sense we have in mind, is essentially compositional.

Linguists sometimes puts the point by saying that an account of meaning must show how the capacity for linguistic competence, at least considered in the abstract, is *generative*, or *productive*. Frege himself stressed that the linguistic capacity is a *creative* capacity.

A theory of meaning, then, must take the following form: first, it assigns meanings to the simple (non-composite) expressions of the language (of which there will be only finitely many); second, it states principles that determine, for a given grammatically correct way of combining meaningful expressions, the meaning of the composite expression.

A WRONG PATH

Following this lead, we might suppose that for each simple or complex expression *e* of the object language, the theory will entail a theorem which says:

The meaning of *e* is ...

where the place of '...' is taken by a singular term that names an object – i.e. a meaning. This seems perfectly natural. The theory would take the form of a *recursive assignment of meanings to expressions*.

However, it will not quite work that way. For what are these meanings? According to Frege, we can distinguish sense from reference. Senses are special sorts of entity. The sense, for example, of 'Theaetatus' is such an entity. But what is this entity? We cannot literally say 'The sense of "Theaetetus" is the man who...', where 'the man who...' is some description such as might be taken to *express* the sense of the name. For such a sentence says that the sense of the name is a certain *man* (the man who ...), which is absurd. All we can really say, by way of specifying the sense of 'Theaetatus', is that it is the sense of 'Theaetatus'. So we are told: the sense of 'Theaetatus' is the sense of 'Theaetatus'. This is completely uninformative; it is merely an instance of the law of identity, i.e. that $x = x$, for any x .

In addition, an explicitly Fregean approach will leave us saying something like 'The sense of "Theaetatus flies" is composed of the sense of "Theaetatus" and the sense of "flies"'. What is meant here by 'composed'? Clearly, *this does not actually tell us the meaning of the sentence, because it does not enable us to understand it*: merely being told what entities go to make up the meaning of a sentence is not sufficient for being told the meaning of the sentence. A theory of meaning requires more than a *list* of entities; we need to know the way the entities go together. Especially striking is the case of relational sentences such as 'Socrates taught Plato'; to be told that the meaning comprises the meanings of the three expressions does not distinguish the meaning of 'Socrates taught Plato' from that of 'Plato taught Socrates'. This is familiar from Chapter 4.

Another way to see the hopelessness of this approach is to observe that, for example, a monolingual French speaker ought to be able to gain explicit knowledge of their language by having a theory of meaning *expressed* in French that is *about French itself*. And if such a theory were translated into English, it ought to tell the English speaker about the semantics of French – in exactly the same way that knowledge of the theory of general relativity can be translated from French into English. But this is blocked under the present proposal: an English translation of the French theory about French will include such sentences as 'The sense of "boire" is the sense of "boire"', which does not tell an English-only speaker the meaning of the French word.

● DAVIDSON'S WAY: T-SENTENCES

Davidson's way forward is to *reject* the idea that the meaning of a sentence should itself be an entity of some kind. That is, we should stop looking for a singular term to fill the blank of

The meaning of e is ____.

Instead, as a first step towards Davidson's actual proposal, consider the following:

s means that *p*

Earlier, we assumed that 'that *p*' is a singular term referring to the proposition expressed by '*p*'. But we do not have to look at it that way. We could think of 'means that' as an expression that joins a name of a sentence with a sentence to form a new sentence. Thus if *s* is a sentence of the object language under study, then the place of '*p*' will be taken by a sentence of the metalanguage, the language we are using to describe that language. An example might be

'La neige est blanche' means that snow is white.

if the object language were French.

Remember the task was to make a theory that would tell us, for an arbitrary sentence of the language, what it means. Would a theory that entailed theorems of the above form do the trick? It looks as if it would, but a very large problem remains for the attempt to devise such a theory. The problem is that the theory would be *using* the expression 'means that'. If we have a true sentence of the form '*s* means that *p*', the result of replacing *p* by another sentence *q* of the same truth-value need not result in a sentence '*s* means that *q*' with the same truth-value as '*s* means that *p*'. In a word, the place occupied by '*p*' in such sentences is *non-extensional* (in fact, in the language of Chapter 11, it is not intensional but *hyper-intensional*). It is unlike the case, say, of '*p* and *q*', where we can replace the sentence *q* with any sentence *r* of the same truth-value as *q*, and the result will have the same truth-value as '*p* and *q*'. In particular, '*s* means that *p*' can be replaced by '*s* means that *q*' only if *p* and *q* have the same meaning. Therefore it is unclear how our knowledge of meaning could be explained in terms of a theory of meaning that uses precisely the form of words 'means that'. *It seems that we would be explaining the concept of meaning only by presupposing the concept of meaning.*

What we really want, according to Davidson, is a theory of meaning that avoids notions such as 'means that'. We want a theory which tells us something *like*

'La neige est blanche' means that snow is white.

– that is, we want it to yield statements which somehow enable us to interpret a *named* sentence of the object language by means of a sentence of the metalanguage, a sentence that is *used*, as 'snow is white' is used in the above. But we want to do it in a way that is purely *extensional*, so that we clearly understand the logic of what we are doing.

Davidson's proposal is that we can do this by using the predicate 'is true'. Consider the following:

'La neige est blanche' is true iff snow is white.

Such a sentence, following Alfred Tarski, is called a **T-sentence**. Generalised, the suggestion now is that the theory should entail a T-sentence of the form

s is true iff p

for each sentence s of the object language (and where ' p ' is replaced by a sentence of the metalanguage). The connective 'iff' – 'if and only if' – is certainly extensional, so its logic and meaning are clear (it has a complete truth-table). Furthermore, the predicate ' α is true' is extensional: from ' s is true' and ' $s = s^*$ ', we can correctly infer ' s^* is true'.

A T-sentence can plausibly be regarded as stating the *truth-condition* of a sentence. But can a T-sentence 'plausibly be regarded' as stating the sense or meaning of a sentence? Does it literally tell us the meaning of a sentence, as ' s means that p ' does? No, it does not in itself. In fact, Davidson's claim is *not* that T-sentences directly state the meaning of each sentence of the object language. His claim is that a theory of truth for the language as a *whole* can be regarded as a theory of meaning for the language *as a whole*. This is Davidson's famous doctrine of **holism**. In order to see the basis of this claim, we need to look in more detail at (1) the structure of a theory of truth and (2) the way in which a theory of truth for a language can be known to be correct.

● AN EXAMPLE OF A THEORY OF MEANING

Since there are infinitely many well-formed sentences of a language, there are infinitely many T-sentences for that language. So if a theory of truth is to serve as a theory of meaning, and a theory of meaning must be something knowable, the theory must admit of being formulated as a finite set of statements. A theory of meaning for an extant natural language such as French is out of the question here; a natural language, even the rudiments of one, is forbiddingly complex and would distract us from our immediate concerns. Here is an example of a theory of truth for a simple 'toy' language L, consisting only of atomic sentences and truth-functional combinations of those, treating only of its typographical aspect, not specifying how it would be pronounced. (The terms ' S_1 ' and ' S_2 ' are used below in the metalanguage as metalinguistic variables; an object is said to 'satisfy' a predicate just in case the predicate is *true-of* the object.)

I Syntax

L consists of:

- a Names: a, b .
- b Predicates: $F\alpha, R\alpha\beta$.
- c Connectives: \forall, ∇ .
- d Sentences: the result of replacing all the Greek letters in an L-predicate with names is an atomic sentence of L. Every atomic sentence of L is a sentence

of L. The result of prefixing a sentence of L with '∇' is also a sentence of L. The result joining two sentences of L with '♥' is also a sentence of L. Nothing else is a sentence of L.

II Semantics

- a The referent of 'a' = Socrates.
The referent of 'b' = Plato.
- b For any x, x satisfies 'Fa' iff x is a philosopher.
For any x and y, x and y satisfy 'Rαβ' iff x taught y.
- c Any atomic sentence of L is true iff the referent(s) of its name(s) satisfies (satisfy) its predicate.
- d If S₁ and S₂ are any sentences of L, then:
'∇S₁' is true iff S₁ is not true.
'S₁ ♥ S₂' is B true iff S₁ is true and S₂ is true.

That's it.¹ Now let us see an example of how the theory entails T-sentences. Consider the L-sentence '∇Fb'. From clause IId, we have:

- (1) '∇Fb' is true iff 'Fb' is not true.

By clause IIc, we have:

- (2) 'Fb' is true iff the referent of 'b' satisfies 'Fa'.

Since clause IIa tells us that referent of 'b' = Plato, we have from (2):

- (3) 'Fb' is true iff Plato satisfies 'Fa'.

So by clause IIb,

- (4) 'Fb' is true iff Plato is a philosopher.

Which is logically equivalent to

- (5) 'Fb' is not true iff Plato is not a philosopher.

So from (1) and (5), we derive:

- (6) '∇Fb' is true iff Plato is not a philosopher.

(6) is a T-sentence for the L-sentence '∇Fb'. Obviously, '∇' is a sign of negation (and '♥' is a sign of conjunction). '∇Fb' is presumably a false sentence, but that does not stop (6) from being true. (1)–(6) constitute a derivation of a T-sentence *for* the language L, but the derivation is carried out in the metalanguage; in particular we use English to state and reason about the semantics of L. (1)–(6) *mention* sentences of L but do not *use* them. (6), a T-sentence for L, is of special interest: *it states the truth-condition of an L-sentence*. (6), along with the whole theory of meaning we have

just given for L, states something that a speaker of L would be expected (implicitly) to *know*: the L-speaker may not speak English, but they do (implicitly) know what is stated by (6), just as they may know, without speaking English, the fact that is stated by 'The moon orbits the Earth'.

In similar fashion, our little theory of truth entails a T-sentence for every sentence of L, even though L, modest as it is, has infinitely many sentences. For more complicated languages, the truth-theory must be more complicated, but our example gets across the basic idea: a truth-theory assigns semantic properties (we can think of 'is satisfied by' as an analogue for predicates of reference) and uses 'recursive' (repeatedly applicable clauses like IIc and IId) to assign truth-conditions to sentences on the basis of those.

● FURTHER DISCUSSION: QUANTIFIERS AND VARIABLES

To make our little toy example into a non-trivial and more realistic pretend language, we would need to add more predicates, singular terms and sentential connectives and, more momentously, we would need to add variables and quantifiers. This would enable one to express, in L, the equivalent of 'All philosophers who have studied with Plato are wise'. Let us see very briefly how this would go.

For ease of reading, we shall now write '¬' for '∇', and '&' for '♥'. (d) is replaced with a definition of 'formula of L' as follows.

We add to the syntax: (a*) Variables $x, y, z, x', \dots, x'' \dots$ and so on, as many as needed. The result of replacing all the Greek letters in an L-predicate with names or variables is a formula of L. The result of prefixing a formula of L with '¬' is also a formula of L. The result of joining two formulas of L with '&' is also a formula of L. We shall take the additional one-place and two-place predicates as understood. Quantifiers: if Φ is any formula of L, then ' $\forall v(\Phi)$ ' and ' $\exists v(\Phi)$ ' are formulas of L. ('For all v', 'For some v'; we use parentheses here in an intuitive manner.) We use 'v' as a dummy letter, standing for an arbitrary variable. v is *bound* iff it occurs in the scope of a quantifier ' $\forall v \dots$ ' or ' $\exists v \dots$ ' (intuitively, scope is found by counting; beginning immediately after the quantifier, the extent of the scope of the quantifier is reached when the number of right brackets equals the number of left brackets). v is *free* just case it is not bound.

Nothing else is a formula of L. A formula with no occurrences of free variables is a *closed* formula. Otherwise it is an *open* formula.

For the semantics of L, we require the notion of an assignment of objects to variables. By an assignment we mean an arbitrary pairing of each object (each object 'in the domain') with a variable in some enumeration of the variables. Each variable is thus like a temporary name. Then, in line with the semantical clauses (a), (b) and (c) – with the adjustment putting '¬' for '∇', and '&' for '♥' – for each such assignment A, every simple formula (formula with no quantifiers or truth-functions but

possibly with variables) will have a truth-value with respect to A . Accordingly, add to the semantic clauses (a)–(d) the following: ' $\forall v(\Phi)$ ' is true with respect to the assignment A iff Φ is true for every assignment which is just like A except at most what it assigns to v . ' $\exists v(\Phi)$ ' is true with respect to the assignment A iff Φ is true for some assignment which is just like A except at most what it assigns to v . A formula is true *simpliciter* iff it is true under every assignment.

Thus consider an L-equivalent sentence of the English sentence 'All philosophers who have studied with Plato are wise'. For ease of reading, assume the truth-function 'if-then' is also added, notated ' \rightarrow ' [it is explicitly definable in terms we are already using: ' $p \rightarrow q$ ' is ' $\neg(p \ \& \ \neg q)$ ']. Thus we can write: $\forall x(Fx \ \& \ Pxb \rightarrow Wx)$. (With the predicate letters as obvious.) This is true iff for every assignment it is true; take an arbitrary assignment. It is true with respect to such an assignment A just in case ' $(Fx \ \& \ Pxb \rightarrow Wx)$ ' is true for every assignment just like A except at most what assigns to ' x '. Which is true iff no arbitrary object both is a philosopher and was taught by Plato but is not wise.

● THE EMPIRICAL CONFIRMATION OF A THEORY OF MEANING: RADICAL INTERPRETATION

Our overarching task was to show how a theory of meaning can be empirically applicable in the third-person style. We've explained that such a theory is really an extensional theory of the truth-conditions of a language and shown essentially how that works via a simple example. But we haven't seen how to devise and confirm such a theory; we've merely looked at a made-up one, true by virtue of stipulation.

So imagine you are a linguist exploring the Indonesian jungle, searching for previously unknown languages. You discover a previously unknown tribe whose language has never previously been identified or even heard by the outside world. Your task, as a **radical interpreter**, is to devise a correct Davidsonian theory of meaning for this language (a theory of truth for it). How do you do it? How do you begin, and how do you know whether or not the theory you come up with is correct (how it is empirically confirmed)?

You begin by hanging out with natives. What sorts of things do you watch out for? Suppose you notice that from time to time they make a certain sequence of sounds, 'Gav-a-gai' (the example is adapted from Quine). Looking out for what they could be talking about, you notice that they say it only when a rabbit appears. Next time a rabbit scampers by, you try saying to one of the natives, in a questioning tone of voice, 'Gavagai?', without any idea of what sounds correlate to the individual words of the language. They smile and say 'Jai!' Now you try it when no rabbit is present: 'Nie!', they say.

You conjecture that 'Jai' is a sign of assent (like 'yes') and 'Nie' a sign of dissent. You now write down, as a provisional hypothesis, that the following is a T-sentence for the language which you dub 'Nove':

'Gavagai' is true-in-Nove iff a rabbit is present.

You do the same with other sentences. So far, you are assigning truth-conditions directly to sentences. But, as you know, this piecemeal, sentence-by-sentence approach could never suffice for the semantics of the whole language. What you need is to assign meanings to sub-sentential expressions, with an eye towards formulating a compositional semantics, a truth-theory for the whole language.

How can you formulate hypotheses concerning sub-sentential expressions? Suppose you come across a sentence whose provisional T-sentence you formulate thus:

'Bavagai' is true-in-Nove iff a monkey is present.

Now you have a clue: surely 'agai' is an expression meaning something like 'a ... is present' or 'there is a ...', and 'Gav' means *rabbit*, and 'Bav' means *monkey*. Suppose by similar means you come to interpret 'Bollo' as *snake*, and try saying 'Bollo-agai' when a snake is present; assent from the natives will tend to confirm these hypotheses.

In these and similar ways you work up a semantics for Nove. Now suppose that one day an *ape* – an orangutang – enters the scene. A member of the tribe shouts 'Bav-agai!' You now have two choices: you could suppose that 'Bav' means not *monkey* but *monkey-or-ape*; this would be plausible, since monkeys and apes are so similar. Or you could suppose that 'Bav' means monkey, but that the member falsely believes that the orangutang is a monkey.

How do you decide which course to take? The two options, let us suppose, look *equally plausible*. You decide, let us say, on one of the two options. Proceeding on that basis, you might find that things run pretty smoothly. But eventually other, similar cases will arise, forcing you to choose between either revising an earlier interpretation or attributing a false belief to the native speaker. As interpreter, it seems your position is this: given such a case, there will always be more than one way to accommodate the observations you make, and accommodating the data in a given way might require you to revise earlier theoretical choices. According to Davidson, there is not going to be a rule for deciding each case *individually*; what you have to do, rather, is to look at the overall patterns you come up with that would accommodate the data and decide which one, if any, fits it best.

This is Davidson's doctrine of *holism*, or rather *meaning holism*, or *semantical holism*. The interpreter's task cannot be described as finding empirical confirmation for the interpretation of each *word* individually or each *sentence* individually but finding the best *overall* fit between theory and evidence. Strictly speaking, it makes no sense to ask, 'is this interpretation of this word correct?'; we can only ask whether an entire theory of meaning is correct, is empirically confirmed. Of course, we do often ask such questions about particular words, but the fact that we can regard an answer to such a question as justified shows that we are always assuming a background interpretation of the language as a whole.

Lurking in the foregoing considerations is a more unsettling idea. It might be that whole patterns of attributions fit the data equally well; they look equally plausible.

In that case, you'll have multiple finished theories that seem individually to be empirically confirmed: you can use them to understand the natives, and you can use them to say things to them. Yet the two theories will disagree; they will depict the native language and beliefs differently. Indeed, if you recognise this possibility, then even if you had just one empirically corroborated theory, you'll know that you *could have* made different choices along the way and come up with a different theory that fits the data just as well.

Such is the *indeterminacy* of interpretation, which was preceded in a slightly different context by Quine's argument that translation is indeterminate. It is important to recognise that Davidson does not regard this as a limitation to his approach; he does not accept that his approach *falls short* of meaning, as if facts about meaning somehow elude the attempt of a radical interpreter. *There are no such facts outside the ken of a radical interpreter.* If there is one way to interpret a language-user, then there are bound to be other ways. This is fine.

● THE PRINCIPLE OF CHARITY AND THE INTERDEPENDENCE OF BELIEF AND MEANING

So far we have been a little bit vague about what it is to 'confirm' a theory of meaning, about what it is for such a theory to 'fit' the evidence and what exactly the evidence *is*.

For Davidson, the principal evidence available to the radical interpreter will be manifestations of the native's *holding-true* a certain sentence in a certain circumstance. For example, the native manifests this attitude by *assenting* to the sentence as described above or by asserting it (just as the interpreter must begin by guessing as to what manifests assent, they must also guess as to what kind of *utterances* of a sentence constitutes *asserting* it).

Roughly, what the interpreter is first looking for are correlations between the attitude of holding-true a sentence and observable circumstances. The interpreter finds, for example, a correlation between holding-true a sentence and the circumstance that it is raining, and infers that the sentence is true if and only if it is raining.

What is most useful, then, are sentences whose truth-values *vary* depending on circumstance, such as 'Gavagai' or a sentence that means that it's raining. These are called **occasion sentences**. Their opposite, **standing sentences**, are not as useful; since natives will always hold-true a sentence that means that granite is harder than wood, and granite always is harder than wood, we cannot find correlations between varying dispositions to assent to the sentence and varying truth-values of the sentence.

In general, the attitude of holding-true a certain sentence is the product of two factors: the native *belief* and the *meaning* of the sentence (Figure 10.1).

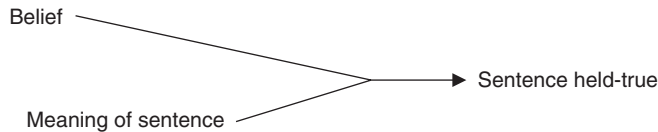


Figure 10.1 The two factors responsible for holding a sentence true

The natives hold 'Gavagai' true in a certain circumstance because they believe that a rabbit is there, and 'Gavagai' is true if and only if a rabbit is present. But once in a while, the natives might have a false belief, thereby holding-true something that's false. Hence what Davidson calls the 'interdependence of belief and meaning'; the role of the concept of belief is just to take up the inevitable slack between *holding-true* and *is true*. If it is false that *p*, but we wish to interpret a sentence that is held-true by a certain subject as meaning that *p*, then we may do so, provided that a *false belief* that *p* is ascribed to the subject.

And now it might occur to you: how do we know that this doesn't go on all the time? After all, we *could* make the interpretation of 'Gavagai' consistent with the data by assuming that the natives are always wrong about rabbits: perhaps they believe that a rabbit is present if and only if a rabbit is *not* present, and we should interpret 'Gavagai' as 'No rabbit is present'. Or we could interpret 'Gavagai' as true if and only if the moon is on fire and suppose them to believe that the moon is on fire just when a rabbit is present. These hypotheses seem bizarre, but what rules them out?

What rules them out is that if such hypotheses are entertained, then *no* theory of meaning can be justified – or, rather, *any* theory of meaning, any scheme of interpretation, could be made to accord with any set of evidence. In order to get started with an interpretation, and in order to sustain it, we *must* assume that the natives are usually correct, especially about relatively obvious things such as whether or not it is raining, whether or not that's a rabbit and so on. This is not like a bootstraps-style 'refutation' of scepticism, where one simply declares the thing done. Think of what goes on when a child learns a language. If the people around the child were in the habit of speaking falsely as often as truly, the child could never guess the meanings of words, since there would be no observable correlation between what is said and what the child observes.

This is Davidson's celebrated **principle of charity**. Since no interpretation – hence no understanding – is possible except under the assumption that speakers generally have true beliefs, the discovery that speakers generally have false beliefs is impossible. In order to make such a discovery, we would have to interpret their language, which requires the assumption of true beliefs. Thus it is a methodological imperative that the subject being interpreted be assumed to have mostly true beliefs. In order to understand another creature, we must attribute beliefs to that creature that are mostly true; we must, that is, assume agreement between ourselves and the subject (compare this with Grice's maxims of cooperation: Chapter 9, 'Implicature').

Is there a hint of *transcendental argument* here? At one time, Davidson drew the more radical metaphysical conclusion that no language-using creature *could* be

systematically wrong (so *our* beliefs are mostly true, so the sceptic is defeated!). For consider the conceivable possibility of an 'omniscient interpreter', the *OI*. The *OI* would be able to interpret any creature (otherwise the *OI* would not be omniscient). But since the *OI*'s beliefs are all true, and interpretation requires the principle of charity, the subject *OI* must attribute mostly true beliefs to the subject. Since the *OI* is omniscient, the *OI* cannot be wrong to do so. Therefore the subject's beliefs are mostly true.

This argument has been challenged. One might think that the *OI* would not have to use the principle of charity, since the *OI*, being omniscient, already knows what the subject's beliefs are. On the other hand, suppose we restrict the *OI* in such a way that it does not already know the subject's beliefs; it knows, say, only all those publicly observable facts that would be relevant to an ordinary interpreter. If interpretation requires charity, hence agreement between interpreter and subject, then perhaps the *OI* *could not* interpret the subject.

● FURTHER DISCUSSION: EVENT SEMANTICS AND INDIRECT SPEECH

A likely thought is that Davidson's programme seems to involve the not inconsiderable assumption that natural language – with all its formidable grammatical complexity – must admit to being represented in the relatively austere grammar of formal logic, to make possible the formulation of a workable theory of truth for the language. We should require indexicals – they are straightforward – but also prepositions, case, tense, relative clauses, auxiliary verbs and more will have to be tamed. In this short section we will consider work in this direction that has been done by Davidson himself: on something known as 'event semantics', and on indirect speech.

Event semantics/adverbs. Compare the following (imagine these are actual speech-acts; the subject is the same Jeff throughout):

- (1) Jeff danced.
- (2) Jeff danced badly.
- (3) Jeff danced badly, doing the funky chicken.

Surely, as a matter of logic, (1) follows from (2), which in turn follows from (3). But how, exactly? As a matter of predicates, (1)–(2) for example seem to be representable only as 'Fa' and 'Ga'. As a matter of the logic of truth-functions, (1)–(2) do not show any pertinent structure. And surely there are no quantifiers.

Or are there not? Concentrate on (2). The adverb 'badly' seems to modify 'danced'. In fact, if you consider all the people dancing at given interval, then a certain subset of them will be dancing badly. The extension of the former contains the extension of the latter. Still no sign of logic, since for example the pizzas consumed by me in the past week are a subset of the items purchased by me in the past week, but that is not a matter of logic. But the idea that we are talking about 'people dancing' furnishes a clue: we are talking thus about activities taking place in time and in space—more

generally, we're talking about **events**. This is not explicit in (1)–(3), but, as Davidson sees it, these sentences all have a tacit quantifier ranging over events. Making the quantifiers explicit, we write for (1) and (2):

- (1*) $\exists e_i[\text{dancing}(\text{Jeff}, e_i)]$
 (2*) $\exists e_i[\text{dancing}(\text{Jeff}, e_i) \ \& \ \text{badly}(e_i)]$

The subscript ' i ' attached to ' e ' is an indexical representing the time and place. Thus (1*) says there was, in a certain place at a certain time, a dancing-event with Jeff the agent of the dance. (2*) says the same but adds that the event was done badly. Clearly (2*) implies (1*). (3) goes as:

- (3*) $\exists e_i[\text{dancing}(\text{Jeff}, e_i) \ \& \ \text{badly}(e_i) \ \& \ \text{the funky chicken}(e_i)]$

Clearly, although 'doing the funky chicken' looks like a sentential clause, it is here being treated as an adverbial phrase; so long as the analysis issuing in (2*) is roughly on the right lines, it seems fully justified to extend it to (3); for (3) does seem to imply (2), and it is demonstrable that (3*) implies (2*).

The account can be broadened to all sentences which can be seen as referring to events (and states), and does address a legitimate semantical need, that of accounting for adverbs and related phenomena. However, it does come at a price, an 'ontological' price: we thereby accept that there are events (and states, which we can think of as the temporary possession of properties).

Parataxis and indirect speech. Especially as readers who have skipped ahead to Chapter 11 will appreciate, but readers of Chapters 2 or 3 will have at least some awareness of, sentences like the following can be devils to settle on an interpretation of:

- (4) Galileo said that the earth moves.
 (5) Galileo believes that the earth moves.

Following Frege, you might take the structure of these to be simple. (5) will be 'Believes(Galileo, that the earth moves)', like 'Kicked(Pele, the ball)'. The trouble, at least from Davidson's standpoint, is that a Fregean is thinking of 'that the earth moves' as referring to a proposition, which is the meaning of a sentence, an intensional entity. (4) will be 'Said(Galileo, that the earth moves)', which is same except that the 'said' relation requires Galileo to have verbally expressed the proposition, in some manner or other. But with our semanticist hats on, to speak of propositions is to presuppose the sharpness and determinacy of meaning, which is of dubious justification; in fact it is a main conclusion of Davidson's approach that it is *indeterminate*, thanks to semantic holism.²

Davidson proposes to analyse (4) in the following way. As a halfway house, we write:

- (4*) Galileo said that. The earth moves.

Which we represent as:

(6) *Samesaid* (Galileo, that). The earth moves.

– where a given speaker x bears the *Samesaid* relation to an utterance u just in case some utterance on the part of x is synonymous with u , and 'that' is a demonstrative which refers to the following utterance of 'The earth moves'. Thus Davidson's 'paratactic' approach to indirect speech ('parataxis' means 'arranged side by side'; in this case meaning that the two sentences about one another). (4*) is obtained from (4) by positing a suppressed full stop (a period) and ensuing capitalisation after 'that'. If we think of ' x samesaid u ' as abbreviating ' x uttered something synonymous with u ', then we can rewrite (2) like this:

(7) $\exists x(\text{Uttered}(\text{Galileo}, x) \ \& \ \text{Synonymous}(x, u)) \ \& \ [u = \langle \text{The earth moves} \rangle]$.

where ' u ' is a demonstrative, referring to an actual utterance, as here indicated by the special brackets (which you can read as 'the utterance of'). Thus (6) is emphatically not a conjunction: the role of the second sentence is only to provide a suitable referent for the demonstrative, and the truth of the attribution depends entirely upon the truth of the first sentence.

The position occupied by 'that' is extensional. Contra Frege, and thankfully for the semantical theorist, the words comprised by the content clauses in indirect speech retain their ordinary meanings. The notion of '*Synonymous*' is as loose and sloppy as you please; it is in effect a feature of the object language used by ordinary speakers, not a general, official concept used in the metalanguage by the semanticist. And the proposal rightly counts as invalid such an argument as 'John said that Marian Evans is wise; Marian Evans = George Eliot; therefore John said that George Eliot is wise'.

As for (5), we should start with:

(8) Galileo believed that. The earth moves.

Again, the parataxis. Then we should require, in place of the notion of 'samesaying', an analogous notion such as 'believes the content of' (condensed as '*Bel_{DD}*'):

(9) *Bel_{DD}* (Galileo, that). The earth moves.

Then we would have the simpler:

(10) *Bel_{DD}* (Galileo, u) & [$u = \langle \text{The earth moves} \rangle$].

• HISTORICAL NOTES

Donald Davidson stresses in his intellectual biography that he was impressed early on by the idea that whatever meaning is, it has to be something public; it has to be

something which, in principle, can be made available to an interpreter, or a child learning language. On this, as well as on some other points, he agrees with Quine (the subject of Chapter 6). He was thanked by Quine in the Acknowledgements of *Word and Object* of 1960, and the two maintained a close working relationship through the 1990s. Davidson thought a great deal before publishing, but once he came out with 'Theories of Meaning and Learnable Language' in 1965 (in Davidson 1984) and 'Truth and Meaning' in 1967 (see 'Primary reading', below), a series of papers came thick and fast, articulating the ramifications and applications of his view and revising it as well. Almost immediately he attracted attention; Richard Rorty (1931–2007), Tyler Burge, John McDowell and Ernie Lepore are some well-known commentators. Davidson's theory of language has been extraordinarily fertile, setting philosophers and linguists to work on finding ways to interpret within a Davidsonian theory of truth-conditions such recalcitrant beasts as adverbs and the propositional attitudes as we've seen, plus predicative adjectives, indexicals and demonstratives, tense, mood, causal statements and more (Davidson himself provided some of this work). He is equally prominent in the philosophy of mind, developing his view known as 'anomalist monism', according to which mental description is not reducible to physical description but nevertheless concerns the same (physical) objects. In later work, he was less sanguine about theories of meaning; since everyone's language differs marginally from everyone else's, and everyone's language is always changing, and everyone occasionally 'mis-speaks' (is guilty of *malapropisms*), those formal objects called 'Theories of Meaning' have to be only idealisations, and as 'passing theories' have to be applicable only in the moment. Indeed, he said that strictly 'There is no such thing as language', if language is the subject of his own theories of meaning.

● CHAPTER SUMMARY

Davidson approaches the phenomenon of meaning with two key requirements: a theory of meaning for a particular language *L* must, for each sentence of *L*, entail a theorem that in some sense states or displays its meaning; and we must explain how such a theory is empirically confirmed. If we succeed in describing such a theory in general – irrespective of the details of any particular *L* – we will have described meaning in general. Since the range of possible sentences constructible in a language is potentially infinite, such a theory must be compositional.

It is not possible, according to Davidson, for the theory simply to take the form of assigning *entities* to expressions, such as Fregean senses. Such a scheme inevitably is uninformative as well as incapable of getting around the list problem. Instead, Davidson first considers the idea that a theory of meaning for a language should issue in true statements of the form '*s* means that *p*' for every sentence *s* of the language. The problem, however, is that such statements are hyper-intensional: they involve the term 'means that' and thus cannot be used in theory whose purpose is to illuminate meaning, without generating a circle. Davidson's proposal is that a theory of *truth* can be regarded as a theory of meaning: the aim should be to generate a true 'T-sentence' of the form '*s* is true if and only if *p*', where *s* is any sentence of the object language and *p* is a sentence of the metalanguage. Tarski had already shown

how to construct such a theory, one that generates an infinitude of T-sentences from a finite base, consisting of reference-assignments for simple singular terms, satisfaction conditions for simple predicates and so on.

The idea is not that a true T-sentence literally states the meaning of a sentence individually: it is that a whole theory that generates all and only true T-sentences conveys all there is to convey about the meanings of the sentences of the language; a theory of meaning is *holistic*, not atomic.

To explain how such a theory is confirmed, it is sufficient to explain how to devise such a theory for a language that is new to us; such is Davidsonian radical interpretation. This will draw out, will make explicit, anything upon which our understanding of a language depends. Occasion sentences – like ‘It’s raining’, as opposed to standing sentences like ‘Wood burns’ – afford the interpreter a way to get started. The native linguistic behaviour – for example, “‘Gavagai’ is true-in-Nove iff a rabbit is present’ – might be tried as a possible T-sentence. Subsequent observations might bear out this hypothesis along with many more. Chopping sentences into parts that are assigned referential and satisfaction conditions will suggest other T-sentences, which can then be tested and so on, until we have a complete theory.

If a native assents to ‘Gavagai’ in the presence of a rabbit, what is presumed to be going on is that ‘Gavagai’ is indeed true if and only if a rabbit is present, *and* the native *believes* that a rabbit is present. In other words, ‘Gavagai’ is *held-true* by the native in a certain circumstance, and ‘Gavagai’ is true in that circumstance. The assumption is Davidson’s principle of charity: without this assumption, interpretation is impossible. Sometimes, however, the pattern is broken: we mistakenly hold a certain sentence true which actually is false. Since we are human beings, interpretation must allow for this. If, once in a while, the native is found to hold-true, for example, ‘Gavagai’, when it is not the case that a rabbit is present, we can stick to our interpretation as ‘A rabbit is present’. But this can happen only within limits; if it happens too frequently, we can no longer pay for an interpretation in the currency of ascription of false belief; we must revise the interpretation. Such is the interpenetration of belief and meaning.

• STUDY QUESTIONS

- 1 People often believe all sorts of wild things, but still they speak and are understood. Is that not in tension with the principle of charity, which requires that the beliefs ascribed be mostly true? Can a more detailed principle of charity cope better? Perhaps something like: assume that the native speakers have generally true beliefs about *what is evident in the immediate environment*?
- 2 If we always had true beliefs, and never had false beliefs, then evidently Davidson would not have to introduce the concept of belief into his theory at all: ‘is true’ would take the place of ‘is held true’. It would appear, then, that there are only *beliefs* if there are *false beliefs*. Is that plausible?
- 3 Holism entails that the meanings of all the expressions of a language are interdependent. Is it therefore impossible to have a language consisting of just one sentence with a particular meaning?

- 4 Holism entails that the meanings of all the expressions of a language are interdependent. Therefore it is not possible to learn the meaning of a single expression without learning the meanings of all. But in that case, it is impossible to learn such a language; one doesn't learn a language at one fell swoop! So holism is false. What can a defender of holism say in response?
- 5 What can Davidson say to someone who advances the following objection? Consider: "Gavagai" is true if and only if a rabbit is present and all kangaroos are kangaroos'. This statement assigns exactly the same truth-conditions to 'Gavagai' as "Gavagai" is true if and only if a rabbit is present'. By your lights, Mr Davidson, they mean the same. But obviously they don't!
- 6 If we imagine a repetition of (6) – call it (6*) – then a larger speech-act comprising (6) and (6*), though it would not entail that Galileo uttered two things, would entail that Galileo *said* two things; whereas the conjunction of (4) with itself evidently does not. Is this objection fatal?

● PRIMARY READING

'Truth and Meaning' and 'Radical Interpretation' are essential; they are both in Davidson's *Inquiries into Truth and Interpretation* (1984) and *The Essential Davidson* (2006). The former includes much more on the philosophy of language, including 'Belief and the Basis of Meaning', 'Reply to Foster', 'Reality Without Reference' and 'The Inscrutability of Reference'. The latter includes 'A Nice Derangement of Epitaphs' and 'The logical form of action sentences'. Both volumes have 'On Saying That' and 'What Metaphors Mean'.

● SECONDARY READING

Evnine, S. (1991) *Donald Davidson*.

Ernest L. and Kirk L. (2005) *Donald Davidson: Meaning, Truth, Language, and Reality*.

Malpas, J. 'Donald Davidson', in the *Stanford Encyclopedia of Philosophy*.

● NOTES

- 1 The use of brackets to indicate grouping has been suppressed, and I have glossed over the need to have, at the very least, ordered pairs as the satisfiers of two-place predicates. It would be easy to add more names and predicates to the language. All truth-functional connectives can be defined using the two given, so nothing essential would be added by adding more of those. Significant new complexity comes when we add quantifiers to such a language, but the platitudinous character of our simple example remains. For most of these points, see the next section.
- 2 A more decisive problem, but a more difficult one, is this: 'believes that ...', if a normal predicate, must act upon the referents of terms put for '...'. But if that is merely a truth-value; then for example 'J believes that q' would follow from 'J believes that p' and 'if p then q'.

11

• the propositional attitudes

● EXTENSIONALITY REVISITED

The principle of extensionality is closely related to the principle of substitutivity (Chapter 2, 'Substitutivity and extensionality'). It is usefully seen as comprehending three parts:

- (i) For any occurrence of a sentence within a larger sentence, the former can always be replaced by another sentence of the same truth-value, and the truth-value of the larger sentence will be unchanged.
- (ii) For any occurrence of a predicate within a sentence, the predicate can always be replaced by another with the same extension, and the truth-value of the sentence will be unchanged.
- (iii) For any occurrence of a singular term within a sentence, the singular term can always be replaced by another with the same reference, and the truth-value of the sentence will be unchanged.

Extensionality is highly intuitive. Consider (iii): if a sentence says that a certain object satisfies a predicate, and another sentence says that the *same* object satisfies the very same predicate, then surely if one sentence is true, then so is the other. For example:

- (1) Paris is north of Vienna.
- (2) 'Paris' and 'the capital of France' are co-referential.
- (3) The capital of France is north of Vienna.

Whether or not (1) is true, it seems that (2) tells you that (3) will have the same truth-value as (1). Similarly for (ii):

- (4) This goldfish has a pancreas.
- (5) 'α has a pancreas' and 'α has a spleen' are co-extensive.
- (6) This goldfish has a spleen.

Supposing (5) to be true, then, again, the truth-value of (4) must be the same as that of (6) (sameness of truth-value, remember, is not sameness of *meaning*).

But English and natural languages generally seem to contain cases showing non-extensionality or substitution-failure. Simple and obvious cases are counterexamples to (i), involving non-extensional **sentence-connectives** such as *because*. For example:

- (7) The cat meowed because the dog barked.
- (8) The cat meowed because Paris is in France.

Assuming that (7) is true, hence that 'The cat meowed' and 'The dog barked' are true, replacing the true 'The dog barked' with the equally true 'Paris is in France' yields (8), which (we assume) is false.

● REFERENTIAL OPACITY AND FREGE ON THE ATTITUDES

Consider:

- (9) 'Venus' has five letters.

The two names 'Venus' and 'the Morning Star' refer to the same object, the planet Venus. But we cannot validly infer:

- (10) 'The Morning Star' has five letters.

Unlike (9), (10) is false. An explanation is that the singular term appearing at the beginning of (9) is not a singular term referring to Venus; rather, it refers to 'Venus' – it's a name of a name, not of a planet. "'Venus'" refers to 'Venus', not Venus. The position in which it occurs is **referentially opaque** as opposed to **referentially transparent**. Inserted into that position, the name 'Venus' does not do its customary job of referring to Venus. That is why the context

- (11) '___' has five letters.

is not open to substitution of co-extensive expressions; it is non-extensional. 'Venus', as it occurs in (9), is simply not in the business of naming Venus. Instead, what happens when we insert an expression into the blank of (11) is that the quotation marks together with the expression form a name of that expression, and the sentence says something about it.

But this context, subtly different from (11), is referentially transparent:

(12) ____ has five letters.

or:

(13) α has five letters.

Frege held that something very similar goes for *indirect* quotation and for the related expressions of propositional attitude, such as belief. Let us briefly revisit the view we introduced in Chapter 2. Consider the following examples, involving indirect quotation:

(14) Adam said that Venus is a planet.

(15) Adam did not say that the Morning Star is a planet.

These can both be true, even though, of course,

(16) The Morning Star = Venus.

For if Adam does not know of the truth of (16) – for example, if he had never heard of *the Morning Star* – then surely both (14) and (15) can be true. Indeed, even if Adam *does* know (16), it seems that the fact that he spoke as (14) says he did does not establish that he actually *said* that the Morning Star is a planet.

Such non-extensional contexts are **hyper-intensional** contexts. Cases like this are also hyper-intensional:

(17) Adam believes that Venus is a planet.

(18) Adam does not believe that the Morning Star is a planet.

(17) is called a hyper-intensional context and not a mere intensional context, because although ‘The Morning Star = Venus’ is a necessary truth (use ‘Venus = Hesperus’ if you like), one cannot reliably substitute the one for the other in (17) whilst preserving truth-value. Unlike the situation of (14) and (15) however, we must, in order to preserve the rationality of Adam, assume that Adam does not know that (16) is true. With that assumption made, it does appear that (17) and (18) could be jointly true. Frege’s explanation is that

(19) that Venus is a planet

is not logically a sentence, but a complex singular term denoting a proposition, the sense of ‘Venus is a planet’. The word ‘that’ converts a sentence into a singular term referring to the sense of the attached sentence. This sense is different from the proposition denoted by

(20) that the Morning Star is a planet

This is reflected in the fact that 'Venus is a planet' and 'The Morning Star is a planet' express different propositions, different senses. It thus appears that the context

(21) Adam believes that α is a planet.

is referentially opaque; the truth of a whole sentence formed by inserting a singular term into (21) does not depend on what is referred to by the singular term. For Frege, it depends on the sense of the term. More generally, the context

(22) Adam believes that S

where S represents any sentence, is referentially opaque; it presents a hyper-intensional context. The general point is that

(23) α believes that S

is referentially opaque with respect to the sentences substituted for S . As with quotation, however, the subtly different expression

(24) α believes β

is simply a two-place predicate, standing for a normal, extensional relation: belief (compare (23) and (24) with (11) and (12), or with (11) and (13)).

Thus for Frege, the opaque context (23) is really very much like (11). A sentence formed by filling the blank of (11) with a referring term does not thereby yield a sentence about the referent of that referring term but about the term itself. Likewise in the case of (23), except that instead of getting a sentence about the *term* inserted there in (11), we get a sentence about two things, namely the referent of the term put for α and a *proposition*, the *sense* of a sentence.

Moreover, according to Frege, a sentence like (17) does *not* contain an expression referring to the planet Venus, or to the Morning Star. The sentence refers to Adam and to the proposition expressed by 'Venus is a planet', and it says that the former believes the latter, but there is no reference to Venus in the singular term 'that Venus is a planet'.

In fact, this is in accord with intuition, because one can assert such things as

(25) Le Verrier believed that Vulcan is smaller than Mercury.

This is true, despite the nonexistence of Vulcan. Hence the truth-value of a sentence containing the belief-operator does not require that singular terms within the scope of the belief-operator refer to their customary objects. All that is required is

that the that-phrases – the indirect clauses as grammarians know them – pick out a proposition. Which is the case in the above example: although ‘Vulcan is smaller than Mercury’ lacks a truth-value owing to reference failure on the part of ‘Vulcan’, it does according to Frege express a sense, a proposition, and the statement says that Le Verrier believed it.

● FURTHER DISCUSSION: MULTIPLE HYPER-INTENSIONAL EMBEDDING

Imagine a cat upon a mat; it is a state of affairs that we’ll call CM. Now consider a picture of CM from a certain point of view: the picture depicts CM in a certain way (the ‘way’ is analogous to the sense of a proposition). Call this picture P1. Now consider P2, a picture of P1, a picture of a picture of CM. P2 depicts P1 in a certain way, which, in turn, depicts CM in a certain way. This can go on – to P3, P4 and so on, as CM is overlaid by multiple ways of depiction – but in principle the information contained in P1 is never lost (we have to imagine ideal pictures). From any picture in the chain, we can recover the information in P1, the state of affairs CM.

The structure of this situation is very much how Frege conceived of propositions about propositions, or propositions about propositions about propositions and so on. For Frege, the contribution of a given term to the determination of a sentence’s truth-value is affected by whether the sentence embeds the term in intensional or hyper-intensional operators. In extensional contexts, the relevant entity is the term’s ordinary referent; in hyper-intensional contexts, the relevant entity is what Frege calls the term’s ‘indirect’ referent, which is the term’s ordinary sense. Speaking loosely, the reference of a term ‘shifts’, and can shift yet higher, as in:

(26) Dudley believes that Adam believes that Don Ho is a singer.

If we call cases like (17) cases of *type-1 hyper-intensional embedding*, involving the term’s indirect sense and its indirect reference, then (26) exemplifies *type-2 hyper-intensional embedding*, involving *doubly* indirect sense and reference.

Table 11.1 may help. The example is the singular term ‘Don Ho’. For ease of reading the elbow brackets indicate *the sense of ‘...’* – for example, <Don Ho> = the sense of ‘Don Ho’, and <<Don Ho>> = the sense of “the sense of ‘Don Ho’”. Don Ho is thus determined by <Don Ho>, which is determined by <<Don Ho>> (which is determined by <<<Don Ho>>>); the black arrows indicate that it’s the same entity, repeated.

The first column depicts the normal, extensional case: in ‘Don Ho is a singer’, the term ‘Don Ho’ expresses its customary sense <Don Ho>, which determines its customary referent, the Hawaiian Don Ho (who was indeed a singer; you may have heard his big hit *Tiny Bubbles*). The second column depicts a type-1 hyper-intensional context – such as that of ‘Adam believes that Don Ho is a singer’. The customary *sense* of ‘Don Ho’ is the term’s indirect *referent*. The third column

Table 11.1 The sense and reference of a term as used in three contexts

Extensional context	Type-1 hyper-intensional context	Type-2 hyper-intensional context
Referent (customary) <i>Don Ho</i>		
Sense (customary) < <i>Don Ho</i> >	→ Referent (indirect) < <i>Don Ho</i> >	
	Sense (indirect) << <i>Don Ho</i> >>	→ Referent (doubly indirect) << <i>Don Ho</i> >>
		Sense (doubly indirect) <<< <i>Don Ho</i> >>>

shows, as in (26), that the indirect sense of the term is the term's doubly indirect referent <<Don Ho>> (so the doubly indirect sense is <<<Don Ho>>>). In principle, the pattern must continue, since there is no end to the possible embeddings of a term using 'believes that'; we have 'A believes that B believes that C believes that ...' and so on. Again, it's useful to think in terms of pictures, to think of a picture of a picture of ... of a subject (which is like the ordinary referent).

• DE RE AND DE DICTO NECESSITY

An important sort of non-extensional context is generated by modal adverbs such as 'necessarily'. We've touched on the topic before (Chapter 7, 'Necessity, possibility and possible worlds: a primer'; Chapter 8, 'Putnam on natural kind terms and essence'). As usual, we assume the customary view that the truths of mathematics are necessary truths, truths that could not have been otherwise. Thus consider:

(27) Necessarily, two is less than three.

That's true. Now Mars has two moons, Phobos and Deimos. So the singular term 'the number of Martian moons' has the same reference as 'two'. But we cannot infer from that fact and (27):

(28) Necessarily, the number of Martian moons is less than three.

It's a fact but only a contingent fact that Mars has less than three moons; Mars might have had more. 'Necessarily', then, creates a non-extensional context. But this is an intensional context, not a *hyper*-intensional context. The reason for this terminological distinction is that modal adverbs do not cut things as finely as propositional attitude verbs do (for example). For example, it is a necessary truth that

any kangaroo is a kangaroo. Therefore the possible worlds in which Beckett was born in 1906 are exactly the same as those in which Beckett was born in 1906 and any kangaroo is a kangaroo. But a person can wonder whether Beckett was born in 1906 without wondering whether Beckett was born in 1906 and any kangaroo is a kangaroo. The semantical difference is not fully reflected in the sentences' intensional truth-conditions, but in their hyper-intensional truth-conditions (which is close, if not identical to their senses, in Frege's terminology).

Sentence (28) is false. But now consider:

(29) Two is such that necessarily it is less than three.

Quine – subject of Chapter 6 – called the transition from (27) to (29) a step of *exportation*. The pronoun 'it' has its reference supplied by 'Two', but, unlike 'Two', it is within the scope of the intensional adverb 'necessarily'. (27) and (29) illustrate the distinction between *de dicto* statements of necessity – of *things said* – like (27), and *de re* statements of necessity – of *things* – like (29).

In (29), the term 'Two' is *outside the scope* of the necessity operator. 'Two' has shifted from an intensional position in (27) to an extensional one in (29). Thus, since two = the number of Martian moons, we *can* infer from (29):

(30) The number of Martian moons is such that necessarily it is less than three.

Like (29), and unlike (28), this sentence is true: it says, consider the number of Martian moons – two, e.g. the number of ears on King Charles, call it what you will – *that object* is necessarily less than three. Exportation slices through the conceptual content of the term to the naked object.

● **DE RE AND DE DICTO BELIEF**

The same distinction can be made with respect to hyper-intensional contexts. To explain and illustrate, let us consider Quine's famous example involving statements of propositional attitude. Ralph has seen a certain man in a brown hat behaving suspiciously. Abbreviating 'the man in the brown hat' as 'the MBH', we write:

(31) Ralph believes that the MBH is a spy.

One might think that this establishes that

(32) The MBH is such that Ralph believes that he is a spy.

It looks just like the inference from (27) to (29): it involves moving 'the MBH' from inside the scope of 'believes' to outside its scope, from a *de dicto* statement of belief to a *de re*. This seemingly innocuous transition from the *de dicto* to the *de re* turns out

to have immense epistemological significance. To bring it out, consider an existential quantification of the *de re* case (32):

(33) There is an *x* such that (Ralph believes that *x* is a spy).

The sentence (33) represents a much more interesting state for Ralph than simply the *de dicto*

(34) Ralph believes that there is an *x* such that (*x* is a spy).

Ralph believes there are spies but so does everyone; the point of (33) is that, unlike most of us, he suspects someone in particular of being a spy.

Considered in itself, the step from (32) to (33) is plainly valid, since the place quantified into – the occurrence in (32) of ‘The MBH’ – is a perfectly normal occurrence of a singular term, an occurrence that is referentially transparent. It is *not* bound by or within the scope of ‘believes that’, but stands outside the occurrence of ‘believes that’ (the second occurrence of the variable ‘*x*’ has replaced the word ‘he’; in this use of the word ‘he’, it has the same function as the variable).

But looking through our Fregean microscope, the inference from (31) to (32) seems not to be valid. For at the end of the section ‘Referential opacity and Frege on the attitudes’, above, we saw that within belief contexts, terms refer not to their customary referents but to their customary senses. (31) states a relation of Ralph to a proposition, *not* to such objects as the MBH. In fact, (31) does not even entail that there *is* such a man, as we learned from thinking about La Verrier and Vulcan; therefore it cannot entail (32), which *does* entail the existence of the MBH. (33), and hence (32), is made true partly by the existence of a certain man, namely the MBH, when strictly speaking (31) makes no mention of that man, the MBH. A referentially opaque context in a sentence like (31) appears to be sealed off from such devices as quantifiers; one cannot, without further ado, *quantify into* positions *within* that context by means of a quantifier placed outside the context. (31) can usefully be compared to a statement about a painting: Ralph’s taking a certain painting to be of real events does not entail that the events actually exist or that elements of the painting themselves refer.

It seems that the further premise that we are assuming in moving from (31) to (32) is that ‘The MBH’ expresses an object determining sense – that is, that it refers. Roughly, the assumption is:

There is an *x* such that: the sense of ‘The MBH’ determines *x*.

● RALPH’S PREDICAMENT

Quine adds to the story of Ralph. A certain Bernard J. Ortcutt is known to Ralph as an upstanding member of society, certainly no spy. Hence

(35) Ralph believes that Ortcutt is not a spy.

Just as we inferred (32) from (31) on the grounds that the MBH exists, and hence (33) from (32), we can from (35) infer:

(36) There is an x such that (Ralph believes that x is not a spy).

on the grounds that Ortcutt exists. Yet, unknown to Ralph,

(37) The MBH = Ortcutt.

Now (33) says that Ralph believes, of someone, that he is a spy. This man is the MBH, otherwise known as Ortcutt. Thus, concerning Ortcutt, Ralph believes him to be a spy. But, by exactly parallel reasoning, concerning Ortcutt, Ralph believes him *not* to be a spy. What is Ralph's attitude towards this man? For Ralph, of course, it depends on how he is thinking of that man; thinking of him as the MBH, he thinks him to be a spy; thinking of him as Ortcutt, he thinks him not to be a spy. But what about for *us*? Obviously, Ralph's attitude towards that man does not depend on how *we* are thinking of that man. Concerning Ortcutt, does Ralph believe him to be a spy, or not?

Having appreciated Frege's lesson of the Morning Star, it is plausible to say: both. We are not charging Ralph with inconsistency or irrationality, as if we were saying of a proposition P that Ralph believes that P and that Ralph believes that not- P . The idea is that in (31), the term we employ inside the scope of 'believes' must reflect Ralph's way of thinking of Ortcutt, i.e. he is thinking of him *as the MBH* – likewise in (35), except that he is thinking of him *as Ortcutt*. In (31) and (35), the position occupied by 'the MBH' or 'Ortcutt' is referentially opaque, not open to substitution by co-referentials. The context 'Ralph believes that ___ is a spy' depends for its applicability on something other than simply the object referred to by whatever singular term we insert into the blank. It is not, as we might put it, simply about the object.

In (32), by contrast, 'the MBH' need not perform any such function: the position occupied by 'the MBH' is fully and transparently referential, open to substitution by co-referring singular terms (such as 'Ortcutt'). The operation of exportation on (31) to yield (32) is a transition from a *de dicto* construction to a *de re* one; (33), meanwhile, explicitly purports to tell us that *there is* a certain object that satisfies that predicate.

The epistemological significance of the *de re/de dicto* distinction is manifold, but there are two phenomena that stand out.

(i) **The indispensably of the *de re***. Russell once suggested the example 'I thought your yacht was longer than it is'. Let A be the yacht in question. Obviously, what Russell is saying about his own beliefs at a certain time in the past cannot be represented like this:

(38) Bertrand believes that A is longer than A .

He does not profess to have believed an explicit contradiction. What we want to say is not that Bertrand thought that the length of A is longer than the length of A, but, rather, that *there is* a length, the actual length of A, such that Bertrand *believes* that the length of A is greater than *it*. To represent this, we must export material from inside the scope of 'believes that' to outside its scope:

- (39) There is an x, and there is a y, such that (x = the length of A & y is greater than x & Bertrand believes that y is the length of A).

That is, in order to describe Bertrand's mistake, we have to quantify from outside into the scope of 'believes'; the *de re* approach is mandatory.

(ii) **A further restriction on exportation.** Suppose Ralph's friend Leo, unlike Ralph, has no reason whatsoever to suspect anyone in particular of being a spy. Still, just on general grounds, he has reason to believe that the two most vertically challenged spies are not exactly the same height. Suppose he's right on that score. So, apparently, we must accept:

- (40) Leo believes that the shortest spy is a spy.

Since the shortest spy exists, we may apply exportation to (40), to yield:

- (41) The shortest spy is such that Leo believes that he is a spy.

And then we may quantify on 'the shortest spy' in (41):

- (42) There is an x such that (Leo believes that x is a spy).

But something has gone wrong. (42) appears to have Leo suspecting someone of being a spy, which we explicitly said is not the case. Again, the inference from (41) to (42) is watertight: it's merely of the form 'B is thus and so' to 'There is something that is thus and so'.

The problem must be with the exportation step from (40) to (41). Merely having a *de dicto* belief appears not to be sufficient for having the corresponding *de re* belief; belief that a proposition containing a referring, descriptive singular term is true is not sufficient for having a belief *about the thing* referred to by the singular term. It is not sufficient for having a belief with respect to a singular proposition involving that object. So exportation – the transition from the *de dicto* to the *de re* – requires something more than we have so far recognised.

It's natural to say that the problem with Leo and the shortest spy is that Leo doesn't *know who* the shortest spy is; if a premise were added to the effect that he does know who the shortest spy is, then surely (41) would be forthcoming. That thought has been famously sharpened, refined and elaborated on by David Kaplan. We'll just say enough to get a taste. Kaplan puts the thought by saying that the inference requires the term subjected to exportation – 'the shortest spy' in (40) – should put the subject of the belief – Leo – into a position of being epistemically *en rapport* with the object of belief (the object to which the *de re* statement purports to relate the believer). The key is Kaplan's notion of a 'vivid designator' – a singular term that is an element of a person's inner story; it is a 'conglomeration' of mental images,

partial descriptions and ordinary names, which, *if* the object exists, serves to bring the object to a given person's mind; and even if the object does not exist, it is from the subject's point of view *as if* it did. Vivacity is what is known in the philosophy of mind as an *internal* phenomenon: a vivid name need not have a referent, and one may be mistaken as to the truth-value of identity statements, even when both sides of the statement are vivid.

The crucial thing is that if a given person thinks in terms of a vivid designator, *and* if there exists an object such that the designator is suitably related to the object for the believer, then, as Kaplan puts it, the designator *represents* the object *to* the believer. In such a case, exportation is allowed. In the non-vivid case – terms such as 'the shortest spy' – the term expresses a certain conceptual halo which might be sufficient to contribute to certain beliefs but which do not sustain exportation. The analogy with Kripke's notion of a rigid designator is intentional.

● BELIEF ATTRIBUTIONS AND EXPLICIT INDEXICALS; BELIEF *DE SE*

Suppose, at a cocktail party, Jones gestures towards Brown, who is across the room, and indulges in the following bit of gossip:

(43) The president of the university believes that he's a charlatan.

How are we to interpret Jones's utterance? On Frege's scheme, strictly speaking, we require the sense that *the president* uses to pick out Brown, *not* the sense Jones uses to pick out Brown. Yet (43) gives us no clue as to what sense that is. Moreover, it seems plain – if uses of indexicals do express senses – that 'he' expresses the sense employed by Jones rather than the president's employment of a sense to pick out Jones.

The problem is exacerbated if Jones generalises his remark about the infamous Brown:

(44) Everyone in the university believes that he is a charlatan.

Jones surely uses 'he' to refer to Brown, and it seems that a Fregean must agree that the indexical expresses the sense that Jones expresses, not somehow all the various senses employed by members of the university to pick out Brown.

In Chapter 8, 'The indispensability of indexicals', we pointed out that indexicals are not just referring devices; they also can serve to locate *oneself* in space and time in a peculiar and ineliminable way. Looking at a map, one sometimes wants to know where *here* is; looking at a calendar, one sometimes wants to know which day it is *now*, or which day is *today*. Mere descriptions couched in terms of general concepts can never quite scratch that itch. Such was the lesson of what Perry calls the essential indexical. The matter crystallises around the use of first-person pronouns – 'I',

‘me’, ‘myself’. For example, if I find during a performance of classical music that the F’s phone is ringing, it is a minor irritation compared with my panic and embarrassment if I learn that the F is *me*. This issue crops up in such a sentence as the following. Imagine, as before, that Jones is the speaker:

(45) The president of the university believes that I am a charlatan.

Jones is now complaining that the president takes a dim view of Jones himself, not of Brown. Again, the indexical pronoun ‘I’ cannot express the *president’s* sense in which *he* thinks of Jones. ‘I’ expresses Jones’ own way of thinking of Jones, of himself. It *seems* therefore that the corresponding proposition – what Jones expresses by ‘I am a charlatan’ – is available to, is thinkable by, only Jones himself, not the president.

We could try to analyse the speech act involving (43) in terms of Kaplan’s scheme, using an existential quantification combined with the indexical:

(46) At the envisaged context, the referent of ‘he’ = Brown, and there is a vivid name *y*, which represents Brown to the president, and the president believes the proposition expressed by ‘*y* is a charlatan’.¹

Something along those lines – similarly for (44) and (45).

David Lewis termed beliefs of this kind beliefs *de se* (‘of oneself’), and argued for a way of characterising them involving ‘property objects’, which feature, if tacitly, points in space-time as well as in logical space, rather than senses or propositions. But to consider it will take us too far astray, and there are other viable approaches to the problem that are closer in spirit to the linguistic philosophies we have discussed in this book. In particular we shall over the next two sections investigate one, one which will seem natural to the extent that Kripke’s picture of proper names seemed natural.

● AN IMPLICIT INDEXICAL ELEMENT

We pointed out in Chapter 7 that the direct referential view of proper names – associated with Kripke – amounts to simply denying the considerations that brought Frege to posit sense in the first place. For if proper names are directly referential, if they have denotation but not connotation, then it remains a puzzle how to explain situations like this:

(47) Alice believes that Marilyn Monroe is a famous actress.

(48) Alice believes that Norma Jean Baker is not a famous actress.

(49) Marilyn Monroe = Norma Jean Baker.

For on the direct reference view, the singular terms ‘Marilyn Monroe’ and ‘Norma Jean Baker’ are synonymous; therefore Alice believes a proposition and its negation,

which is apparently irrational on the part of Alice; but surely the situation represented by (47)–(49) can happen without irrationality. Whereas since on Frege's view, 'Marilyn Monroe' and 'Norma Jean Baker' express different senses, no such contradiction is imputed to Alice.

But there is a powerful reason for doubting that the Fregean line is satisfactory. Suppose you hear a little about Paderewski – that he was the second prime minister of Poland.² Then later, you catch a glimpse of a pianist in the film *Moonlight Sonata*, who is also called 'Paderewski' (in fact, he played himself in the film). It never occurs to you that these are the same man; but they are: one man, one name. So it seems that you accept a proposition – that Paderewski is a politician – and its negation – that Paderewski is not a politician. *It's crucial to note that this problem is independent of the thesis of direct reference*: however the term 'Paderewski' is explained – as Fregean or Kripkean – it appears that you accept that Paderewski is a politician and that you accept that Paderewski is not a politician. It looks like you accept a logical contradiction.

Frege does have a sneaky dodge: perhaps, *in your idiolect*, there is not one name 'Paderewski' but two, corresponding to what you think are two men. But the sort of phenomenon observed in the Paderewski tale is capable of occurring almost anywhere. It is usually at least *conceivable*, if not at all likely, that what you had assumed to be a single self-same object or type of object are in fact two or even more than two. Maybe what you thought of as your left hand is in fact a sequence of entities, one for each day, created by evil scientists; you could name them 'my left hand₁', 'my left hand₂', ... Maybe your left hand₁ ≠ your left hand₂. Since you can't *absolutely* rule out such possibilities, though they are vanishingly unlikely, it seems that they are not the sort of things to be decided by semantical decree, by stipulation. They are not *analytic*. And some cases are not so unlikely. Suppose you look out your window and see, from the side, the front part of a limo; then you look through another window on the same wall and you see the back part of a limo. You wonder – quite reasonably given the various lengths of stretch limos – 'Is that limo that limo?'

The singular terms in this case are not proper names or natural kind terms but indexicals; in particular they are different occurrences of the same demonstrative, each with a different accompanying (implicit) demonstration. But we pointed out in Chapter 8 that Putnam quite plausibly thinks that natural kind terms such as 'water' contain a tacit or hidden indexical element. In the nineteenth century what people were calling 'jade' was discovered to be two distinct minerals, now called by mineralogists jadeite and nephrite. Each such case can be made sense of by making explicit the hidden indexical or demonstrative: *this* (pointing at a sample ordinarily called jade) versus *that* (pointing at another).

In view of such cases, it looks as if, were we to hang on to Frege's conception of language, we would have to maintain that what we vaguely call 'English' is really a vast assemblage of idiolects, each exquisitely complex and fine-grained. But to persist in this is to renounce one of the principal aims of Frege's theory: to lay bare the common, public structure underlying the surface hubbub of speech.

● DIRECT REFERENCE, THE ATTITUDES AND THE SEMANTIC *DE RE*

The foregoing paragraphs suggest that the propositional attitudes are not intertwined with semantics in the way one might think upon studying Frege. In fact, 'the Morning Star' and 'the Evening Star' are very far from being typical proper names; only for those cases is it straightforwardly plausible that each, as a matter of public meaning, has associated with it a distinctive *mode of presentation*. And even for those cases, perhaps all that is going on is that certain properties tend to get defeasibly attached to the term, so as to create the illusion that they are attached still more intimately to the term – that they are attached to it as a matter of the very meaning or sense of the term. And cases such as jadeite and nephrite, Paderewski and the limo in the window are rare, and no surprise: language, among other things, serves the purpose of communication, and it would merely impede that purpose if you required a separate name for each object or portion of stuff you've encountered, just on the off chance that they are distinct objects or substances. If it talks like a duck, then label it as a duck.

Suppose, then, we throw Frege aside in favour of direct reference theory. The semantic or informational value of a proper name, natural kind term or indexical is nothing more than its referent; such is Millianism, the view that proper names and indexicals refer without expressing descriptive content (such also is *Russellianism*, for such is what you get if you embrace Russell but retract the view that ordinary proper names are abbreviations of definite descriptions).

Thus Nathan Salmon has proposed the following alternative account of the attitudes. Suppose 'Superman' and 'Clark Kent' are directly referring proper names (so we are assuming, contrary-to-fact, that the story of Superman is a true one). According to Salmon's view, the information-value of the two names, the semantic content of the two names, is identical. Nevertheless, when Lois Lane hears 'Superman is a hero!' how she responds is very different from how she responds when she hears 'Clark Kent is a hero!' In particular Lois *grasps* this proposition in different *guises* – grasped in one guise, she is disposed to assent to 'Superman is a hero', and grasped in another guise, she's disposed to *dissent* from 'Clark Kent is hero', despite that fact that the two statements are statements of the very same proposition.

The key point is that the different guises under which Lois grasps the proposition are not part of the semantics of the names 'Superman' and 'Clark Kent'. The guises are not a part of the meaning of the statements *Superman is a hero* or *Clark Kent is a hero*. Instead, they are cognitive or psychological features *in Lois* that need not show up in an arbitrary person's use of the crucial terms 'Superman' and 'Clark Kent' – not even if the person is fully competent in using the terms. A person who knows all about Clark Kent's identity with Superman might understand the terms as interchangeable (or if that strikes you as wrong, use another example of a person called by different names).

Let us now apply this way of thinking to sentences that explicitly contain propositional attitude terms. The crucial thing is that in Salmon's theory, guises must appear in an analysis of the attitudes, of belief sentences. If Lois, for example, believes that Superman

is a hero, then there is a guise and a proposition such that Lois stands in certain cognitive relation to the proposition under that guise; call this relation 'BEL', which is a three-place relation corresponding to the ordinary two-place belief relation. Generalising, and understanding '*p*' as indicating a proposition and '*g*' as indicating a guise:

A believes p if and only if: there is a *g* such that [A grasps *p* by means of *g* & A stands in the BEL-relation to *p* under *g*]

Suppose as before that 'Superman is a hero' and 'Clark Kent is a hero' express the same proposition – call the proposition 'P' – but are associated by Lois with different guises; then we have the following:

Lois grasps P by means of a guise that incorporates 'Superman is a hero' & Lois stands in the BEL-relation to P under that guise.

Lois grasps not-P by means of a guise that incorporates 'Clark Kent is not a hero' and Lois stands in the BEL-relation to not-P under that guise.

Thus Lois comes out as believing a proposition P and also its negation not-P: she believes that Superman is a hero but also that Superman is not a hero. She would, of course, deny that she has such a belief as the latter, *expressed in those words*. But just because one dissents from a sentence doesn't establish that one doesn't believe the proposition it expresses. She is not thereby characterised as *irrational*; that charge would stick only if she both believed and disbelieved a proposition under the *same guise*, which is not the case. The same goes in other cases. One might reasonably dissent from the sentence 'Mark Twain is Samuel Clemens' yet believe the proposition all the same: the proposition is just of the form $a = a$, but the guise under which one grasps it in denying it is of the form $a = b$.

We saw in Chapter 9 that Grice offered an explanation as to why certain sentences are semantically correct but pragmatically improper, such as 'If Paris is in Germany then Paris is in Morocco': according to the classical truth-functional account of conditionals, that sentence is true, but the statement made by it, normally, is pragmatically improper. Similarly, Lois Lane believes that Superman is Clark Kent, or, in the above situation, that Clark Kent is a hero, but it would be pragmatically improper to say it in those words.

Guises take up the cognitive slack left on the ground by the semantic properties of expressions. A complete theory of guises would be part of a complete theory of cognition, of the processing of information. Salmon, of course, allows that normally we do endeavour, in ascribing propositional attitudes, to convey something of the subject's point of view; we often try to use the same or similar words that the subject would assent to in characterising his or her attitude. Presumably there is a Gricean maxim according to which we ought in general to attempt to convey more about the cognitive states of believers than what is conveyed purely semantically by the words we use. But still this is a matter of what Salmon calls 'pragmatically imparted information'; it remains the case that it is possible to characterise anomalies such as Lois' while maintaining that proper names are directly referential and do not express Fregean senses.

The following schema is *not* therefore *valid* (it is false in some instances):

If B believes that S, then if B is sincere, B would assent to 'S'.

On this way of thinking, the *de re/de dicto* distinction collapses where proper names are involved. On the older, Frege-derived way of thinking, a *de re* statement of belief requires that the term specifying the object of the belief – 'Hesperus' or 'Marilyn Monroe' – must be outside the scope of the belief-operator or predicate; being *de re* was thus thought to be a syntactical or structural matter. However, the recognition of the direct reference of singular terms does not support this; for if a term such as 'Marilyn Monroe' is directly referential, it contributes nothing to any statement beside its referent, in which case there cannot be a difference in content or truth-conditions between 'Alice believes that Marilyn Monroe is a famous actress' and 'Of Marilyn Monroe, Alice believes that she is a famous actress'. We are led to recognise the *semantical de re* as well the syntactical (or formal, structural) type.

● HISTORICAL NOTES

After World War II, interest in the propositional attitudes picked up dramatically, and although it doesn't dominate the airwaves as it did, it remains very much alive. Frege and Russell set the table long before the war, but it was Rudolf Carnap and Alonzo Church, working broadly in a Fregean framework in the late 1940s and the 1950s, who really got the subject up and running. Carnap came out with *Meaning and Necessity* in 1947 (Carnap 1956), then Church with his 'A Formulation of the Logic of Sense and Denotation' of 1951 and 'Intensional Isomorphism and Identity of Belief' of 1954. Then came Quine's 1955 paper 'Quantifiers and Propositional Attitudes' (in Quine 1975), according to which Fregean referential opacity cannot be all there is to the attitudes – the attitude cannot be *merely* a relation between a subject and a Fregean proposition – because quantification into the scope of a propositional attitude verb seems ordinarily to make sense. Quine ultimately decided against ordinary common sense – in fact proposed a scientific surrogate which has attitudes as relations to sentences rather than propositions – but many people took up his challenge, including Kaplan, who published his landmark paper 'Quantifying In' in 1968 (Kaplan 1969). The 1970s saw a move away from Frege and towards Millianism, or direct reference theory (as in Chapter 7); Kaplan moved that way in later papers, and Kripke published 'A Puzzle About Belief' in 1979, in which he makes the case that, even aside from issues over quantification, Fregean semantics is not nearly so good as one might have thought in coping with belief-puzzles. Nathan Salmon published his *Frege's Puzzle* in 1986; it was one of several attempts to combine an approach to the propositional attitudes with direct reference; others in recent years have included Mark Richard's and Scott Soames'. Salmon's approach is amenable to the 'Language of Thought' theory of Jerry Fodor (1975), according to which cognition takes place in a mental, individual language which has in effect a more Frege-like structure, though this will be at the level of idiolect, not public language. Another stream is the 'paratactic' approach of Donald Davidson, according to which sentences ascribing attitudes do not indicate relations between a subject

and a proposition but between a subject and a token of a sentence supplied by the ascriber. See Chapter 10, pp. 188–189.

● CHAPTER SUMMARY

In Frege's scheme, expressions appearing in hyper-intensional contexts do not refer to their customary referents but to their customary senses: 'Venus' in 'Venus is a planet' refers to the planet Venus, but in 'Bob believes that Venus is a planet', 'Venus' refers to the sense of the term 'Venus', not to its referent, not to Venus. This explains why 'Le Verrier believed that Vulcan orbits the sun' can be true although 'Vulcan' has no referent. And it explains the apparent consistency of 'Bob believes that Venus is a planet' and 'Bob does not believe that the Morning Star is a planet'. Hyper-intensional contexts are referentially opaque, meaning that substitution of co-referential expressions within such contexts is blocked. These phenomena can be reiterated endlessly: we have hyper-intensional contexts, hyper-intensional contexts within hyper-intensional contexts, and so on, as in 'Jim believes that Bill believes that Fred believes that p'.

Unlike hyper-intensional contexts, intensional contexts *de dicto* do not require sameness of sense for substitution; they require only the preservation of the modal characteristics of the expression being substituted for. '1 + 1' can be substituted for '2' in the *de dicto* 'Necessarily, $2 < 3$ ' but not the co-referential 'The number of Martian moons'. Full referential transparency is characteristic of *de re* necessity – such as 'The number of Martian moons is such that necessarily it < 3 '; such statements are implied by but do not imply the corresponding *de dicto* statements of necessity. Quine calls the step from the *de dicto* to the *de re* 'exportation'.

At first blush, exportation in propositional attitude sentences also seems valid, so long as the term subject to exportation succeeds in referring. If *a* is a referring singular term, then 'Bob believes that *Fa*' seems to imply '*a* is such that Bob believes it to be *F*'. This seems to hold, even if there is another singular *b* such that $b = a$, and Bob believes that not-*Fb*; for then *b* is such that Bob believes it to be not-*F*, in which case *a* is such that Bob believes it to be not-*F*. A certain object is such that Bob both believes it to be *F* and believes it not to be *F*; no irrationality is ascribed to Bob in the way it would be if, for example, we charged Bob with believing *Fa* and believing not-*Fa*.

However, exportation seems to require more. For sentences like 'Ralph believes that the shortest spy is a spy', even where the shortest spy exists, do not seem to sustain exportation. A common suggestion is that what is missing in such a case is that Ralph must know who the shortest spy is if exportation is to go through. David Kaplan suggests a lucid formalisation of the idea, requiring that the term subjected to exportation be a *vivid designator*; such designators are like rigid designators in that they sustain exportation and the converse, importation, but unlike them in being subjective – i.e. whether or not a designator is vivid potentially varies from person to person.

The broadly Fregean picture explored so far comes under severe strain from considerations involving indexicals and demonstratives. Certain propositional attitude sentences explicitly contain indexicals, as in Jones' statement 'The president of the

university believes that I am a charlatan'. We learned from John Perry that such *de se* statements are not epistemically equivalent to any statement free of the first-person pronoun, yet 'I am a charlatan' as used in the statement does not express a proposition available to the president of the university. Furthermore, it appears that a Frege-type puzzle can be generated for any object of reference – for example, if one speaks somewhat slowly, one could conceivably wonder whether 'That sun = that sun', where the statement is in fact true, and the first occurrence of 'that sun' is accompanied by a different demonstration from the one accompanying the second. It thus appears that the propositional attitudes do not map onto the contours of language as straightforwardly as one may have thought upon reading Frege. The attitudes are potentially much finer, and the cognitive states they involve more various, than what can be straightforwardly expressed in public language.

Nathan Salmon allows the free substitution of co-referring proper names within the scope of operators of propositional attitudes. Since he regards proper names as directly referential, for example 'Hesperus' and 'Phosphorus' are interchangeable in propositional attitude contexts, the account of propositional attitudes becomes part of the philosophy of mind; one does not merely believe a proposition, but one believes it under some guise, where the make-up of such guises is ultimately a matter for the psychology or cognitive science. We do communicate information about guises, which in many cases align with the theory of Fregean senses, with the crucial difference that the information is only pragmatically imparted, implicated in the manner of Grice, and is not part of the semantics of the terms. Since Lois believes that Superman is brave, she therefore believes that Clark Kent is brave, even if she sincerely asserts 'Clark Kent is not brave'.

● STUDY QUESTIONS

1 Consider:

Hob thinks a witch has blighted Bob's mare, and Nob wonders whether she (the same witch) killed Cob's sow.

The sentence can be true, even though there are no witches. How can this be made sense of? How can the apparent connection between Hob and Nob be reconstructed when the thing that would tie them together doesn't exist? (The example is taken from Geach 1967.)

- 2 Could a brain-in-a-vat have precisely the same beliefs as you according to Fregean semantics? Would typical beliefs such as that *the cat is hungry* be false or neither-true-nor-false, or would a brain-in-a-vat be incapable of such beliefs? What about according to direct reference semantics?
- 3 A likely condition for expressions being synonymous – for their expressing the same sense – is that they be interchangeable in *all* contexts (save quotation). But suppose 'to buy' and 'to purchase' are synonymous. Consider:
 - (a) Nobody doubts that whoever believes that x buys a hot dog, believes that x buys a hot dog.

This and the proposed condition of synonymy jointly entail that

- (b) Nobody doubts that whoever believes that *x* buys a hot dog, believes that *x* purchases a hot dog.

Is that correct? Does (b) express *exactly the same* proposition as (a)? (The example is taken from Mates 1952.)

4 In this book, we have been blithely assuming that, for example,

- (a) that the cat is white

means

- (b) the sense of 'the cat is white'

But surely, *translation* is a matter of preserving sense. A French translation of (a) would be:

- (c) que le chat est blanc

and a French translation of (b) would be

- (d) le sens du «the cat is white»

For (b) pertains to a particular English sentence; it contains a mode of presentation of that sentence. Therefore, since sense determines reference, a translation must preserve reference to that sentence. So (d) is synonymous with (b) but not with (c). Since (c) is synonymous with (a), (b) is not synonymous with (a). But if (b) is not strictly equivalent to (a), what is? Nothing? Or is there some flaw in the argument we've just been through? (Church 1950)

5 Are the guises in Salmon's theory just Fregean senses by another name?

● PRIMARY READING

Kripke, S. (1979) 'A Puzzle about Belief', in A. Margalit (ed.) *Meaning in Use*.

Salmon, N. (1986, 1991) *Frege's Puzzle*.

More difficult:

Kaplan, D. (1969) 'Quantifying In'.

Quine, W. V. (1975) 'Quantifiers and Propositional Attitudes', revised edition.

● SECONDARY READING

Lewis, D. (1983 [1979]) 'Attitudes De Dicto and De Se'.

Forbes, G. 'Frege's Problem: Referential Opacity', *Internet Encyclopedia of Philosophy*.

● NOTES

1 For simplicity I have quantified into the quotation marks, which is strictly speaking nonsense.

2 The example is adapted from Kripke (1979).

12

• chomsky's science of language and Universal Grammar

Historically, philosophers have investigated natural language in terms of the artificial language of formal logic. Frege and Russell were the primary instigators, but one can see it also in more modern figures such as Kaplan and Kripke, Quine and Davidson and, in a less pronounced way, in Grice and Strawson. With this comes a commitment – sometimes merely implicit but often quite explicit as in certain points in Davidson or Frege – that the formal categories of singular term, predicate, quantifiers and so on, and the semantical concepts truth, reference and perhaps sense, are adequate for describing actual human linguistic understanding and the content of arbitrary sentences of natural language.

Frege and Russell both thought of their formal languages as in the first place adequate – adequate for representing the content necessary to codify inference – for logic itself, and then for mathematics, then for the physical sciences. They thought of it as a symbolic notation adequate for performing such services for all fact-stating discourse whatsoever. Frege saw logic as the 'Laws of Truth'.

However, what Frege called a 'Begriffsschrift' – a 'concept-script' – is, in the end, an artificial notation. It was invented, not discovered. Frege made some tremendously fruitful observations relating the structure of natural language to the structure of Begriffsschrift, but those observations, if not merely piecemeal, do not add up to anything like a general analysis of ordinary or natural language. Natural language – even sticking with the fiction that nothing besides English or German need be discussed – is far too nuanced and complex to support the assumption that the techniques and categories of formal logic are adequate to it. Furthermore, in philosophy we tend to go from natural language sentences directly to their formal representations in an intuitive manner, and we tend to make unargued judgements of grammaticality for sentences of natural language.

For illustration consider the following pair:

- (1) Mary expected John to leave.
- (2) Mary persuaded John to leave.

One might initially assume their grammatical structures are the same, thus anticipating a parallel logical analysis. However, a bit more prodding reveals a striking difference which does seem structural. In (1), Mary expected a certain event, namely John's leaving. But (2) does not say that Mary persuaded a certain event. That would be nonsense.

Or consider these:

- (3) John is easy to please.
- (4) John is eager to please.

Again, initially one may expect them to have the same logical analysis. However, (3) entails

- (5) It is easy to please John.

But (4) does not entail

- (6) *It is eager to please John.

As indicated by the asterisk, (6) is nonsensical, or rather ungrammatical (unless 'it' is taken as a referring term, but in (5) 'it' is 'pleonastic' – i.e. is not understood to refer some object). Since what follows from what is *ipso facto* a matter of logic, there is something lacking in the superficial idea that these are of the same logical structure.

These examples, along with many, many others, and the foregoing considerations along with some others which we will discuss below, are what motivate Noam Chomsky in proposing and defending, in a resolutely scientific spirit, the theory of **Universal Grammar**. 'UG', as it is known, aims to state the abstract structure of sentences of natural language directly – and thus ultimately to explain the relevant differences in the way (1) as opposed to (2), and (3) as opposed to (4), are understood. Furthermore, as we will see, it aims to uncover the principles underlying *all* human language, for these have their basis in genetics. In contrast with the very general shortcoming just mentioned of the approach of formal logic-oriented philosophies of language, Chomsky armed with UG is able to formulate much sharper criticisms of certain features of orthodox philosophy of language, such as its reliance on the concept of reference, the emphasis on naming, the assumption that the function of language is communication, the assumption or argument that meaning must be public, and the tendency towards behaviourism. We will discuss all of these, along with just enough of the mechanics or technical details of UG to see something of the basis for the more philosophical conclusions that Chomsky reaches.

● PLATO'S PROBLEM AND BIOLOGY

Chomsky has been enormously productive over nearly seven decades. Chomsky did not invent modern theoretical linguistics, but he has been by far its most influential figure. This is to say nothing, of course, about his activity as a political and social critic; on those fronts he has produced over twenty books as well as being a regular speaker all over the world. Throughout this period, his views in linguistics have undergone substantial changes to be sure, some of which we will discuss very briefly near the end of the chapter. But in the Philosophy of Language and the philosophical end of Linguistics, his basic views have remained largely unchanged.

A foundational point is that he thinks of natural language as a phenomenon to be investigated strictly as an empirical science. This might seem innocuous, but not only did it swing against a large portion of the tide, it is the 'strict' part that has had the most profound repercussions. For what emerges is that human language is not to be studied as a branch of sociology or anthropology, but as a branch of biology.

To see the force of this, consider some observations expressed in Chomsky's semi-historical work *Cartesian Linguistics* (1966).

First, Cartesian dualism has come to have a bad name among many philosophers. But according to Chomsky, its real problem is not that we cannot understand how consciousness could arise in a physical world, so much as that we do not really understand the appellation 'physical' to begin with. In Descartes' time – the early seventeenth century – physical forces were commonly identified with mechanical forces, with causation being explained as involving physical contact, things pushing and pulling. Newton as it were upset the applecart, with gravity entering the picture as accounting for action at a distance – for example, the moon being prevented from flying off into space by its being bound by the earth's force of gravity. Newton himself, despite his theory of gravity's being a decisive advance on previous theories, could not bring himself to believe that it was the last word on the subject. Gravity seemed an 'occult' force, even to him. If we run ahead some three hundred years, the situation has grown decidedly worse for the old view that the physical is to be modelled by colliding balls: space-time relativity, the three subatomic particles, the thirty-six (as of this writing) fundamental particles, the continuity of matter and energy, quantum entanglement, the wave-particle duality of photons etc., dark matter – all we can really say is that physics is concerned with all that, the whole Kahuna, and that we are merely hoping for a unified theory in view of the apparent conflict between Relativity Theory and Quantum Theory. That consciousness or mentality arises within such a world in all its bizarreness must seem not so much of a surprise (indeed, one must positively be prepared for surprises, as anyone who pays attention to headlines regularly coming out of CERN, or the James Webb Telescope, will attest).

Second, Cartesian Rationalism has by no means been refuted. Quite the opposite: even the presumed arch-empiricist David Hume with his 'custom' and 'habit' acknowledged that human beings must have certain innate propensities to belief and inference, in accordance with induction (the inference from the observed to

the unobserved), the existence of the external world, and the continued existence of objects in space, which no amount of reasoning *a priori* can establish. They must be posited, unless one remains a sceptic about those things. An especially pertinent example of a plausibly innate feature of the mind was pointed out by Alex von Humboldt (1769–1859) in what we met with in the Introduction and Chapter 2 as the Principle of **Compositionality**, also known as the Principle of Infinite Generativity: as is evident from the simple example of the multiple iterativity of ‘the mother of x’ – ‘the mother of the mother of x’, ‘the mother of the mother of the mother of x’, and so on – human beings have the capacity to understand sentences and phrases of arbitrary length without upper bound, indeed implicitly to grasp rules which in principle determine the meaning of an infinite number of sentences and phrases, despite their stemming from a finite ‘lexicon’ – that is, a finite base of words and rules.

Third, another historical lesson in favour of rationalism derives from a philosopher of long ago, namely Plato – or, if Plato’s tale is to be believed, from Socrates. In *Meno*, Socrates, and his interlocutor Meno, call in a slave boy. Socrates manages, with a sequence of yes/no questions, to get the boy to demonstrate a theorem of geometry (that the square of twice the area of a given square is a square whose sides are equal to the diagonal of the given square). What the boy exemplifies, according to Plato, is that knowledge is recollection: the boy already ‘knew’ the theorem, but had to be reminded of it by a series of questions. Descartes and Leibniz reached similar conclusions, that knowledge – at least theoretical knowledge – is instilled in the mind by God. ‘How comes it that human beings, whose contacts with the world are brief and personal and limited, are nevertheless able to know as much as they do know?’, asked Bertrand Russell, expressing what is known as ‘Plato’s Problem’ (Russell 1948 p. xxv). The answer was known schematically to Darwin, and not by adverting to the divine.

Plato’s Problem reaches a somewhat more focused form in an argument that Chomsky himself articulated in the 1950s, and which has since been known as the Poverty of the Stimulus argument. A child is not normally ‘taught’ language, let alone subjected to training like a seal. Indeed, for some aspects of what is acquired, no human being had the capacity to teach it until very recently, because the aspects were not even recognised until the 1960s and after. For example, consider the following quartet of sentences (adapted from Chomsky 1986, p. 8):

- (7) John ate an apple.
- (8) John ate.
- (9) John is too stubborn to talk to Bill.
- (10) John is too stubborn to talk to.

One might think that ‘is too stubborn to talk to’ in (9) and (10) plays a precisely analogous role to ‘ate’ in (7) and (8); they are each a predicate, it might seem. (8) means that there is something that John ate, perhaps an apple as is reported by (7). Therefore, by analogy with (7) and (8), since (9) says of John that he is too stubborn to talk to Bill, (10) ought to say that John is so stubborn that he won’t talk

to an arbitrary person. But that is wrong: (10) says that John is so stubborn that an *arbitrary person* won't talk to *him*. The subject of the transitive verb 'talk to' has switched places with its object. (10), certainly, does not say that there is someone that John is too stubborn to talk to. The surface masks what is going under the hood, yet *children do not normally make the relevant mistake, and do not have to be taught this rather abstruse sort of fact*.

In general, the child appears to pick up language based on very limited (and often degraded) experiential information in a very short time. This supports the biological approach. Chomsky concludes that it is more apt to speak of the growth or development of language in the child, much as we speak of the ability to stand, to fight infections or to undergo puberty. More abstractly, Chomsky likens the project of UG to Galileo's early attempt to give a mathematical description of motion and gravity: just as Galileo persevered despite the difficulty of experimentation – for example, the difficulty presented by air-friction to the experimental verification of his hypothesis that things fall at same rate irrespective of their weight – so the difficulty of the project of establishing the principles of UG should not make us give up.

● INTERNAL LANGUAGE AND EXTERNAL LANGUAGE

Substantial features of language are innate or inborn, as aspects of the child's genetic endowment. This is UG, Universal Grammar: the genetically determined component underlying the full gamut of human language. Of course, no one is born speaking English, Portuguese or Japanese; one requires linguistic exposure in order to learn language. To sharpen further the difference between the genetically determined, universal component and the experiential, relatively adventitious component, Chomsky divides one's 'language' into two concepts (assuming for simplicity that one is monolingual): '**Internal Language**' – I-Language – one's individual on-board linguistic competence that is the idiosyncratic joint project of one's genetically determined facility together with the results of one's experience of a particular language; and '**External Language**' – E-Language – which is what is traditionally called 'a language' ('English', 'Portuguese' or 'Japanese'). E-Languages are by no means sharply distinct and are often distinguished largely for political, social or historical reasons (thus the Yiddish saying that a language is just a dialect with an army and a navy). People speaking Dutch can understand people speaking German more readily than people speaking certain different dialects of the self-same Mandarin understand each other, observes Chomsky. To an intelligent extra-terrestrial who does not speak in accord with UG, Chomsky surmises, humans might appear to speak the same language. E-Language is only of peripheral interest for the Chomskian linguist. The Chomskian view is dead set against the idea, associated with Wittgenstein, of the primacy of public language and social institution as opposed to that of the individual capacity, as well as being against the 'structural' style of linguistics which was dominant when Chomsky entered the scene – the theories of Saussure, Bloomfield and Harris that language can profitably be understood as a disembodied corpus of public, interconnected utterances.

● ELEMENTARY PHRASE STRUCTURE GRAMMAR

A key and very general point is that human languages display relations of structural dependence, not of linear or sequential dependence. For example:

(11) Some person took the last one.

We can derive the passive form:

(12) The last one was taken.

Everyone knows – at least implicitly or in practice – that the passive involves (i) deletion of the subject ‘Some person’; (ii) movement of the object ‘the last one’ to the front of the sentence; and (iii) re-writing the main verb ‘took’ as the passive ‘was taken’. But this knowledge is completely independent of the fact that the phrase ‘the last one’ comprises the fourth, fifth and sixth words of the active form. If instead of ‘took’ we had the example ‘took and ate’, we would still be able easily to form the passive, but ‘the last one’ would be the sixth, seventh and eighth words.

Many other types of examples make it still plainer, but the point seems obvious: our linguistic understanding deals in multiple embeddings of phrases – from the inside out – not the absolute sequential ordering of words. To make a passive out of (11), to perform the central operation (ii), we scan the sentence for the direct object phrase – wherever its position in the sentence – and move it to the front. Chomsky stressed the point early, in his *Syntactical Structures* of 1957, and made the point much more rigorously. But the more intuitive understanding is enough: when hearing a sentence, one has to keep track of what has been heard for one cannot always tell the significance and grammatical role of each word until one reaches the end, or at least a point beyond the present word. The same goes for reading a sentence from left to right (or right to left as in Arabic or Hebrew, or top to bottom in traditional East Asian languages). No doubt, the fact that we do hear and read sentences sequentially made this harder to spot.

To introduce some mildly technical vocabulary, ‘Some person’ in (11) is called a *Noun Phrase* or NP, and ‘took the last one’ a *Verb Phrase* or VP. The sentence can be schematised thus:

(13) $S \rightarrow NP VP$

The arrow here means ‘consists of’. This pattern is exemplified by virtually all sentences, in line with the traditional grammar of subject and predicate. The VP itself can be broken down. The VP comprises the *Verb* – V – ‘took’, and another NP, ‘the last one’:

(14) $VP \rightarrow V NP$

In the general case, to indicate the possibility of intransitive verbs – e.g. ‘yawned’ – one could put brackets around ‘NP’ in (14).

Quantifier-words like ‘all’, ‘some’ and ‘the’ in this context are called *determiners*, or ‘det’. Adjectives are treated here as parts of NPs, and are notated ‘Adj’. Thus, with ‘S’ indicating a sentence, we can represent the elementary phrase structure of (11) in terms of nested brackets:

(15) [S[NP[det some][N person]][VP[V took][NP[det the][NP[Adj last [N one]]]]]]]

But many find tree-diagrams easier to read as well as more suggestive of the idea of structure-dependence:

(16)

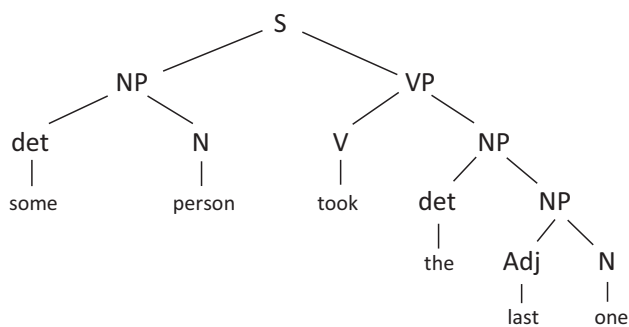


Figure 12.1 ‘Some person took the last one’

Beside nouns and their phrases (N and NP), verbs and their phrases (V and VP), adjectives (Adj) and determiners (det), we also require the category of prepositions and their phrases (P and PP), as in this representation of ‘Sally is from San Francisco’, the preposition being ‘from’:

(17)

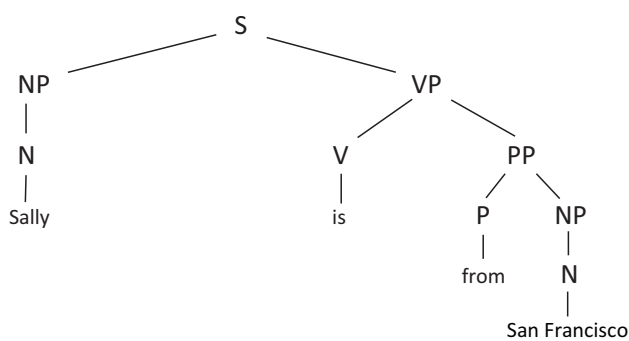


Figure 12.2 ‘Sally is from San Francisco’

● **TRANSFORMATIONAL GRAMMAR; TRACE THEORY**

Consider another example:

(18) The tall girl won the race.

A tree diagram for this sentence would be this:

(19)

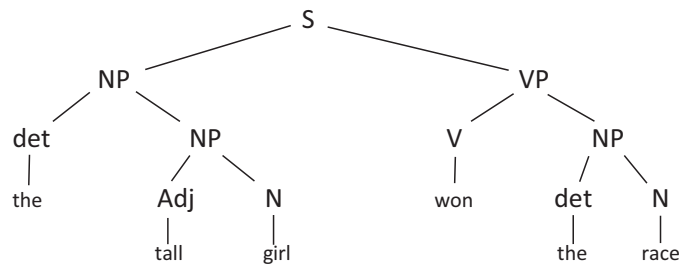


Figure 12.3 'The tall girl won the race' (I)

However, this sentence – as should be familiar for those recalling Russell – should be seen as involving in effect two sentences or clauses: one expressing that a certain girl is tall, and the other that the girl so-described won the race. That is how we understand this entire sentence, but so far the diagram doesn't record this fact. We require something like this:

(20)

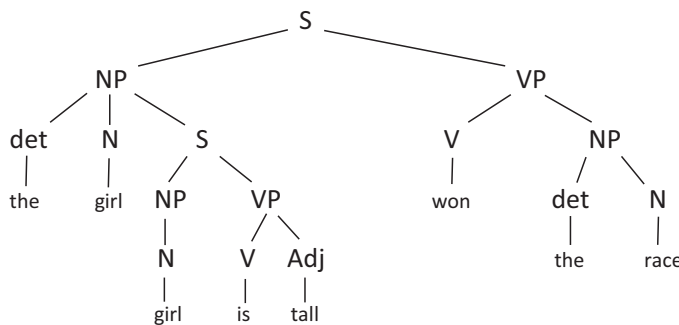


Figure 12.4 'The tall girl won the race' (II)

As a first approximation, (20) represents the 'deep structure' of (19), which for the moment we can think of as representing the 'surface structure' of the ordinary sentence (18). (20) represents two sentences, one inside the other. Then we adopt certain rules to derive the surface structure from the deep structure. Starting with the deep structure, the rules in play, roughly put, are: (a) Replace the most deeply embedded NP with a *wh*-word, in this case the second occurrence of 'girl' by 'who'. (b) Delete 'who is'. (c) Invert the remainder of the VP of the embedded sentence – 'tall' – with the N of the adjoining NP – 'girl'.

This is enough to get across roughly the idea of *Transformational Grammar*: it is the transition from deep structure to surface structure (to prevent unwanted associations, Chomsky came to prefer '**D-structures**' over 'deep structures').

Surface structures (**S-structures**) are also abstract theoretical representations, not captured by mere strings of words that form grammatical sentences. Consider again:

- (3) John is easy to please.
- (4) John is eager to please.

Taken at face value, these ostensibly have the 'same structure'. But as pointed out before, (1) is equivalent to 'It is easy to please John', whereas (2) is not equivalent to the ungrammatical 'It is eager to please John' (with 'it' understood pleonastically – i.e. not as a referring term). This difference is reflected in the derivational history of a surface structure form being displayed using a 'trace', marking the location linked to the moved item:

- (21) John is easy to please *t*.
- (22) John is eager *t* to please.

This indicates which side of 'to please' the sentence places 'John', as subject or object; in the deep structure, this will be explicit.

The same approach can be employed in the case of questions:

- (23) Charlie beat Sam in chess.
- (24) Charlie beat who in chess?
- (25) Who did Charlie beat in chess?

(25) is derived from (24) by moving 'who' to the front, and prefacing the subject with 'did'. This is made more explicit by adding a trace to (25), the position from which the pronoun was moved:

- (26) Who did Charlie beat *t* in chess?

Traces are 'abstract' but must have some sort of mental if unconscious realisation.¹ Obviously, this is just a peek. To make the procedure credibly serviceable and to

see the full range of its power, it must be sharpened, generalised and made more rigorous. And it is merely implicit in the discussion so far that the techniques are applicable to any human language. Nevertheless, we can perhaps see what Chomsky (1975) was getting at when after discussing certain closely related matters, he wrote:

This discussion has been restricted to English, a serious limitation. Nevertheless, I have not hesitated to suggest that the principles that appear to have explanatory power for English are principles of universal grammar. On the assumption that the language faculty is common possession, the inference is plausible (though, obviously, nondemonstrative) ... On the assumption of uniformity of the language capacity across the species, if a general principle is confirmed empirically for a given language and if, furthermore, there is reason to believe that it is not learned (and surely not taught), then it is proper to postulate that the principle belongs to universal grammar, as part of the system of 'pre-existent knowledge' that makes learning possible. (Chomsky 1975, p. 118)

● THE TOTAL LANGUAGE FACULTY

Intended at a very abstract and general level, Chomsky provides the following diagram summing up UG (1986 p. 68):

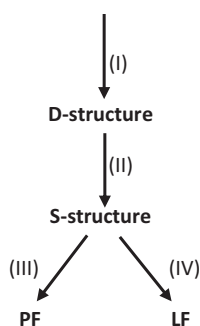


Figure 12.5 The Basic Structure of Universal Grammar (UG)

(I) indicates simple 'Phrase Structure Rules' or 'Base Rules', centred on a finite lexicon that thereby generates an infinite class of D-structures (deep structures). (II) Transformational Rules convert these into S-structures (surface structures). PF is Phonological Form: given an S-structure as input, (III) outputs a phonological item – a phonological 'representation' – which can be used in the making of noises or signs. LF is Logical Form: given an S-structure as input, (IV) outputs a semantic representation (although this is strictly 'internal' semantics, indeed that semantics which can be represented in terms of a broad notion of syntax; more on this presently).

So far, we have spoken somewhat blithely of the procedures of grammar and transformative rules. We will not go beyond an impressionistic account of the procedures,

including certain varieties of what came to be called 'movement'. We shall simply list some of them and make certain comments (some of which we have already seen in action), simply to fill out the diagram just seen, and to make UG out as a plausible account of the human language capacity:

- **X-bar theory.** This goes under (I) at the top of Figure 12.5. It is a theory of how D-structures as it were begin, how they arise from the lexicon; it includes an account of basic phrase structures involving nouns, verbs, prepositions and adjectives. (1980, p. 145; 1986, pp. 80–84)
- **θ-theory** (theta-theory). Enters the diagram at (II) but is also connected with LF (semantics, in one sense). Imposes further constraints such as that the verb 'to borrow' must involve the borrower, the borrowed, and someone from which the borrower borrowed the borrowed. Whereas 'to fall' does not have a subject; one may fall, but one does not fall anything. The potential θ-roles are three in number: Agent, Goal and Patient. Thus the information for 'adore' might be written: 'adore [-- NP]; Agent, Patient'; this indicates that 'adore' goes with a NP, and expresses the agent's being in certain relation to what is referred to by the NP, the 'patient'. (1986, pp. 93–96)
- **Trace theory.** As already demonstrated, letters like 't' are used to indicate positions from which NPs have been moved; that the brain 'keeps track' of such positions in some analogous way is part of the UG hypothesis. (1980, pp. 146–149)
- **Bounding theory** imposes constraints on the distance that phrases can be moved. 'Which pasta did Maria make?' has an S-structure which can be represented along the lines of (18) as '[_S which pasta did [_S Maria made t]]'. 'Which pasta' has been moved from the place indicated by *t*: 'Maria made t', that is 'Maria made which pasta', represents the sentence before the movement. The procedure is grammatically legitimate because the pronominal 'Which pasta' jumps out of just one nesting (one 'bounding node'). But, in English (see below), if we start with 'Stefano didn't realize that Maria made t', by following a similar course we get the ungrammatical 'Which pasta did Stefano didn't realize that Maria made t', represented as '[_S which pasta did [_S Stefano didn't realize that [_S Maria made t]]]'. The reason is that the item represented by the trace jumps too far; the positions must be 'subjacent'. (1986, pp. 71–72, 101–104)
- **Binding theory** deals with the details of the reference of nouns and pronouns, especially anaphorical pronouns, as in 'herself' of 'Sally spilled the wine all over herself', or 'he' in (in a normal reading of) 'If a man fails at this, he ought to be angry' (1986, p. 78). Consider 'He gave Stephano the wine', 'Stephano gave him the wine' and 'Stephano gave himself the wine'. 'He' and 'him' cannot refer to Stephano in the first two, but 'himself' must refer to Stephano in third. Whereas in 'Stephano asked Sally to give him the wine', 'him' can refer to Stephano or some other person. The aim of the theory is to formulate principles governing the sometimes complex referential relationships amongst nouns and pronouns, in a way that interlocks with the rest of UG.

- **Case theory** might seem to have limited application to English, as it deals only with differences between 'she' and 'her', 'he' and 'him', and 'I' and 'me', and the genitive 's' used to create a possessive noun, as in 'Maria's pasta'. The Romance languages such as Italian have the gendered case for all nouns; some languages have more cases – Czech for example has seven: nominative, genitive, dative, vocative, locative, accusative and instrumental. However, in Chomsky-style UG the phenomenon of case exists even where it is invisible – not 'morphologically realized' – but it interacts with other components to yield principled explanations of further varieties of movement between D-structures and S-structures (1986, p. 74).

There are other components such as Inflection and Auxiliaries – together they deal with tense, modality and verb-agreement – but this is enough to get the flavour of UG. The theory is that all these components interlock in such a way as to generate all and only the grammatical sentences of a language, the grammatically significant relations between them, and furthermore a rich characterisation of the meaning of the grammatical sentences, insofar as they can be so generated strictly within syntax.

● KNOWLEDGE, RULES, PRINCIPLES AND PARAMETERS

The foregoing might be described as a complex system of rules; indeed, Chomsky characterised UG thus in the book *Rules and Representations* of 1980. But the idea of 'rules' has the unfortunate connotation of some agent laying down the rules by fiat, with some kind of normative or commanding 'ought' being invoked, of prescription rather than description. Worse, speaking of rules in that sense seems to invite the Wittgensteinian puzzle of rule-following. Chomsky, rather, came to characterise a language-user as acting in accordance with the *principles* of UG, where the principles are wired in to the brain if abstractly (Chomsky 1986, p. 243ff). The principles can be said to be 'cognised', indeed can be said to be 'known' as part of one's 'mental state', but this is allowed only because those terms are ordinary ones, in all their flexibility. There is nothing to suggest that the principles of UG need be available to introspection (let alone conscious or 'access conscious'), any more than the principles according to which one computes the direction of sounds from the forces striking the two eardrums are available to introspection. They are mere facts. The force of calling them 'mental' is simply they are intimately linked with features that are unimpeachably mental, such as thought and perception.

Much the same goes for Chomsky's use of 'representation', as when Chomsky says that a certain D-structure is 'represented' in the mind. It is not meant to suggest an entity standing in a picturing-like relation to some other entity, perhaps witnessed by

an inner homunculus. The mind – or rather the brain – carries out a certain function which corresponds to a certain theoretical description as part of UG, nothing more.

But there is a deeper reason for the shift away from rules. You may be familiar with the fact that whereas in English one says, 'I am studying', in Italian one says 'Sto studiando' – i.e. 'Am studying'. Italian is what is called a 'pro-drop' language, whereas English is a non-pro-drop language (sometimes in casual speech one drops the first-person pronoun, but this is an optional affectation whereas in Italian it is a regular feature). Similarly, other languages *do* allow what we said was ruled out in Bounding Theory (so the item represented by a trace jumps out of more than one nesting). These are not counterexamples to UG but cases where, although no principle is involved, what Chomsky calls a 'parameter' is involved. UG itself, in addition to certain principles which are straightforwardly species-universal, encompasses an array of parameters which have the character of switches, with just two settings, on or off. When first exposed to a language, the parameters are set; only minimal experience is required to trigger them. Another example is English's being a 'Subject-verb-object' or 'SVO' language, whereas other orderings are realised in other languages. Children catch on, setting the switch. Thus the 'Principles and Parameters' approach which Chomsky espoused from c1981: 'The P&P approach held that languages have no rules in anything like familiar sense', he writes, and '[t]here are universal principles and a finite array of options as to how they apply (parameters)' (1995, pp. 5–6). This dovetails with such pronouncements as:

On the one hand, [UG] must be compatible with the diversity of existing (indeed [humanly] possible) grammars. At the same time, it must be sufficiently constrained and restricted in the options it permits so as to account for the fact that each of these grammars develops in the mind on the basis of quite limited evidence. (Chomsky 1981, p. 3)

● MEANING, REFERENCE AND CONCEPTS

The concepts of physics do not correspond neatly to the notions or concepts of common sense. Neither do the concepts of chemistry or biology. We should not therefore expect the concepts of the science of language to correlate smoothly with the concepts of common sense. This comes to a head when we ask about the concept of meaning. This word is commonly used in various ways, covering the relation between a name and its bearer, between a word and its definition, between an individual and an intention, aim or goal. In philosophy, as we have seen, it is used for the relation between a name and its sense, between a general term and the things falling under it, between a general term and its sense, between a sentence and state of affairs, between a sentence and those it implies, between uses of an indexical or pronoun and its referent, and more. Chomsky does talk of 'meaning', of 'semantics', but he intends a particular sharpening of the notion. He does not intend meanings as the entities that are referred to by

words, or meaning-as-the-determinant-of-reference, not if reference is a relation to objects in the environment. He means an *internalist* notion, roughly the product of syntax – using all but only the procedures of UG – together with the concepts investigated by cognitive science (Chomsky 2000, p. 132). Again, it is essentially a concept of hard-core biology, whatever can be explained in terms of biology but only that.

Hilary Putnam, by contrast, put forth the *externalist* idea that ‘meaning ain’t in the head’: that facts of meaning consist at least partly in the context over and above what can affect the creature causally. This is not part of the UG picture. But Chomsky also positively objects to Putnam’s idea, and to related ideas of others, on more specific grounds. Consider ‘water’, which for Putnam is a natural kind term. The average cupful taken from the Tijuana River in Northern Mexico is very impure, that is, contains many other substances besides H₂O. But it is still water. Yet, despite the fact that a cup of tea will contain proportionally more H₂O than the sample from the Tijuana, it is tea, not water. ‘Water’ is not a natural kind term, according to Chomsky; there is certainly no overwhelming reason to go along with Putnam’s intuition that XYZ, a functionally identical substance to H₂O but different chemically, could not be water. Genuine terms for natural kinds tend rather to be artificial terms used in science.

This point of view carries over to the subject of naming. ‘Languages do not seem to have a category of pure names’, avers Chomsky. ‘Rather there are personal names, place-names, color names, and so on’ (1975, p. 45). The reason that Nixon could not have been a pencil, to take one of Kripke’s examples, is not this amazing fact of metaphysics, that a certain homo sapiens considered apart from human ratiocination, could not have been a pencil. It is simply that ‘Nixon’ is a personal proper name. To engage in the practice of calling him by a personal proper name is to presuppose that he was a person (1975, p. 47; we do call pets by personal proper names, but this is best seen as parasitic on their central use). Contrary to the impression generated by referential semanticists, we do not call arbitrary objects – pencils, stones or doorways – by such words as ‘Nixon’. ‘Noting that an entity is named such-and-such’, Chomsky continues, ‘the hearer brings to bear a system of linguistic structure to place the name, and a system of conceptual relations and conditions, along with factual beliefs, to place the thing named. To understand “naming,” we would have to understand these systems’ (1975, p. 46; see also Chomsky 2000, pp. 148–152, 171–173).

Chomsky – after more extended consideration of examples – concludes that

[t]he question, ‘to what does the word X refer?’ has no clear sense ... proper names ... have rich semantic-conceptual properties ... Something is named as person, a river, a city, with the complexity of understanding that goes along with these categories. Language has no logically proper names, stripped of such properties; one must beware of what Peter Strawson called ‘myth of the logically proper name’. (2000, p. 181)

Much the same goes for many general nouns such as 'house'. As Chomsky says, if we are told that a house is painted red, we take its exterior to be painted red, not the interior (2000, pp. 125–126). A house can be knocked down and re-built with new material, or re-built using the same material but after an interval, and be moved from one address to another. These are remarkable entities, but it isn't as if these intricate features are empirical discoveries; it is our concept 'house' at work.

As indicated earlier, for Chomsky, a complete internalist semantical theory would include both UG dealing with the rich varieties of syntax, and a separate theory dealing with concepts – that is, part of the discipline of cognitive science. The possibility that much of its material will turn out to be inborn rather than learned must be reckoned a live one. The poverty of the stimulus bites here as well, if not perhaps so transparently as it does with respect to the acquisition of language.

Chomsky agrees with Strawson that reference is not a property of words or phrases, but is a property of utterances (this orientation should be familiar for readers of Chapter 9). It is a feature of actual speech, of 'speech acts' to invoke Austin or Searle. Strawson, however, defended the view that language is public or social, a view associated with Quine (who famously said that language is a 'social art'), and Wittgenstein (who conceived language as consisting in customs, social practices or institutions). As you might have surmised already, and as we'll see more explicitly in the next section, Chomsky disagrees.

● LANGUAGE, COMMUNICATION, BEHAVIOURISM; PERFORMANCE VS. COMPETENCE

Quine notoriously criticised Chomsky as failing to see the force of his claim, in Quine's terms, that 'translation is indeterminate'. We can understand the claim as being that grammar is indeterminate – that the sum total of physical facts cannot choose between materially different grammars which nevertheless produce the same set of grammatical sentences.

As was mentioned before in connection with (3) and (4), Chomsky does not think the task of UG as merely that of generating the right set of sentences – its task, among other things, is also to explain grammatical relationships between sentences – but here we may set this aside. For although Chomsky admits that the possibility of alternate grammars cannot be ruled out, he does not recognise this so-called criticism as anything but question-begging. Chomsky is not saying the facts he is after are queer, non-physical facts. True, in a certain abstruse sense familiar from the Philosophy of Science, an account of Chomskian UG, or a particular grammar, must be empirically underdetermined – if one account is constructed, then other, empirically equivalent accounts must be possible. But this is true of *all* non-trivial scientific theories. Quine's criticism has bite only if the facts of grammar are ultimately non-physical facts; this is precisely what Chomsky is against.

In *Rules and Representations*, Chomsky writes:

Ultimately, the study of language is a part of human biology. In the study of any organism or machine, we may distinguish between the abstract investigation of the principles by which it operates and the study of the physical realization of the processes and components postulated in the abstract investigation. Thus, the study of visual perception might lead to the hypothetical construction of certain abstract components – for example, feature detectors – that enter into this system. A further inquiry might reveal the physical mechanisms that meet the abstract conditions postulated. In studying some automaton, we might attempt to determine its program at an abstract level, then proceed to inquire into the circuitry or mechanical principles by which this abstract program is realized ... In the study of humans, direct experimental inquiry into physical mechanisms is generally impossible because of the ethics of experimentation ... Therefore, the abstract level of inquiry must bear an inordinate burden. It is important to realize that there is no issue of principle here, no philosophical problem unique to this inquiry as a result of the limitations on feasible experiment. Similar problems would arise in the study of an inorganic device that for some reason we could not take apart. (Chomsky 1980, pp. 226–227)

So if generative linguistics remains mostly at the abstract or functional level, one can still maintain that one's ultimate aim is a description of structures and processes of the brain, of neurology.

It is because of this orientation that Chomsky does not accept behaviourism, or that the relevant evidence for UG is confined to linguistic behaviour (unlike Quine, who said that 'in linguistics, one has no choice' but to accept behaviourism; Quine 1992, p. 37). We have seen that our own judgements or intuitions are appealed to – even if they are not regarded as sacrosanct – but there are other sources of evidence as well: brain-imaging, studies of electrical activity of parts of the brain such as Broca's area, studies of brain injury and disability, analogical reasoning from other species, experiments involving perceived displacement of 'clicks' in speech, suggesting that the subjects really do perceive speech as divided into phrases rather than into sequences of words or phonemes.

It is not surprising, then, to find that Chomsky does not accept that the purpose or function of language is communication, where that means the transmission of information from one individual to another (1975, p. 64). We noted early on that in Chomsky's estimate, children do not learn language by virtue of being taught. It's rather that the faculty of language grows spontaneously in the child, so long as some minimal triggers in the child are activated. Once up and running, the child does communicate of course, but also prattles to itself, reads stories and recites rhymes for its own amusement, later keeps records, lists and diaries, composes more sophisticated poetry perhaps not for the eyes of others, and so on. To say that this is all 'self-communication' is fine but is mere word-making.

For Chomsky, that languages appear for all the world to be public and social can be explained by the 'uniformity of the initial endowment, which permits only

I-languages that are alike in essential respects, thus providing some empirical reason to adopt a version of the Fregean doctrine that “it cannot well be denied that mankind possesses a common treasure of thoughts which is transmitted from generation to generation” (Chomsky 2000, p. 33; the quotation is from Frege 1997, p. 154). Not only do they have the appearance of being public or social, they are, at an essential level, uniform. Though for different reasons, presumably Chomsky would think well of Wittgenstein's famous remark that if ‘lions could talk, we could not understand them’ (Wittgenstein 1958, p. 223): it is a quite different creature – cognitively and genetically different for Chomsky, with different forms of life for Wittgenstein.

Further support for the picture comes from the fact that the project is concerned with linguistic *competence*, not with linguistic *performance*. The latter is sporadic, sloppy and unpredictable, whereas the former is nearly a steady state of the organism, even if accessible only indirectly. If the thesis of UG is correct, then a mature individual's state of linguistic competence is there to be discovered, and changes slowly if at all. In contrast, what the creature actually says and understands depends not only on what is appropriate for the creature in the specific situation and the constellation of the creature's personality, desires and beliefs, but on the creature's *will* – which may, so far as we know or understand, be free (according to Chomsky). Prospects for a ‘theory’ of performance seem, at least at this time, beyond our capacity.

● A RECENT DEVELOPMENT, AND THE POSSIBLE ORIGINS OF LANGUAGE

The highly abstract diagram shown in Figure 12.5 featured two varieties of rules or principles: (I), base-rules that convert items from the lexicon into phrases or sentence-like objects of D-structure, and (II), transformational rules that convert D-structures into S-structures. These rules interlock and are mutually supporting, but still there is no guarantee that none are specific to particular languages. This was partly addressed by the Principles and Parameters approach of the 1980s. In more recent work, Chomsky supplies an alternative and yet tighter architectonic for UG, addressed in turn to (I) and (II).

- Rules under (I) are replaced by a single operation, the ‘Merge’ operation, one which given two objects X and Y, joins them. Set-theoretically, it is the elementary pair set operation: from X and Y, form {X, Y}. X and Y are arbitrary, in the sense that one might be part of the other or might not be. If Y is already a subset of X, then the Merge operation yields an object with a certain hierarchy – perhaps where $X = \{Z, Y\}$, it yields $\{\{Z, Y\}, Y\}$. This is ‘Internal Merge’. If neither Y is a subset of X nor X is a subset of Y, then we have ‘External Merge’. If the entities fed into the operation Merge are word-like entities, then one can see very roughly the basis of the formation of phrases. (Berwick and Chomsky 2016, pp. 72–73)
- Rules under (II), sometimes called the rules expressing the ‘displacement property’, are downgraded in favour of single principle: Move-*a*. This allows us, for any

phrase of any category α , to move it anywhere. This apparently alarming freedom is only apparent: 'How it applies is determined by general principles interacting with the specific parameter choices – switch settings – that determine a particular language'. (Chomsky 2000, p. 13)

One thing that is especially striking about this development – called the 'Minimal Theory' – is that it holds out some hope for the vexed question of the emergence of language. Various other animals – apes especially but a great variety of other species as well – make noises and communicate in at least a rudimentary fashion. What is missing – there are claims that it exists in animals, but let us set this aside – is compositionality, the potential for infinite generativity from a finite lexicon. All humans appear to have it (except for pathological cases), and no non-human animal has it. Merge is a 'recursive' operation, and therefore if it is shown to be integral to the language capacity, it would seem to be sufficient to explain compositionality: from X and Y it outputs {X, Y}; then from X and {X, Y} it outputs {X, {X, Y}}, and so on, without end. If some aspect or bit of the brain could be shown to realise it, this would explain the presence of the crucial property of human language. Furthermore, this historical emergence could in principle be explained by a single genetic mutation. The crucial property of human language thus would *not* have evolved – a word which suggests a gradual process – so much as appeared all at once.

• HISTORICAL NOTES

Chomsky's long and richly productive career can be divided roughly into three periods. First there was his initial splash including the work *Syntactic Structures* (1957), the 'Standard Theory' with the intensive interest in transformational grammar of *Aspects of the Theory Syntax* (1965) and the accompanying philosophical argument of *Cartesian Linguistics* (1968), *Language and Mind* (1972), *Reflections on Language* (1975) and *Rules and Representations* (1980); second was the 'Principles and Parameters' approach of *Lectures on Government and Binding* (1981) and *Knowledge of Language* (1986); and finally the emphasis on Move- α and then Merge, with *New Horizons* (2000), the *Minimalist Program* (2001), and *Why Only Us?* (2016). Except for the book of essays *New Horizons*, this ignores his many papers and some books as well. I have presented UG more or less as a unitary theory, have concentrated on its more philosophical and perhaps more popular aspects, and have barely scraped the surface of its more technical aspects. I have taken the theory post *Aspects* but before the advent of *The Minimalist Program* as representing UG (with Figure 12.5), and then closed by presenting the main post-*Government and Binding* ideas as a later change. The survey is largely impressionistic, but I hope it is enough for the reader to see Chomsky's case for UG as philosophically substantive, formidable and original. As John Collins emphasises, Chomsky's work should be not thought of as scientific with the philosophy an optional add-on; he should be classed with Descartes, Leibniz, and I would add Einstein, all of whom saw their scientific work as necessarily continuous with philosophy (Collins 2008, p. 11f).

A significant proportion of linguists disagree with Chomsky. It is true to say that Chomsky's work in the 1950s and 1960s gave linguistics a new identity – linguistics as branch of cognitive science, which was itself a new discipline, a branch of psychology conceived as part of biology – as well as contributed many of its great technical discoveries. But there have been other ways of approaching the problems, and alternative ones that have arisen since Chomsky's arrival – for example 'Connectionism', which in linguistics involves the attempt to explain the acquisition of language in terms of 'general-domain learning' – as well as 'generative linguistics' not covering all of linguistics. Nor is Chomsky the sole figure behind the UG programme, even if he was the prime instigator. On the linguistic side, one can name Paul Pietroski, Ray Jackendoff, Paul Postal and Jerome Katz; on the philosophical side, one can name James Higginbotham, Peter Ludlow, Georges Rey and John Collins, who was mentioned above.

The seeds of Chomsky's dispute with Quine are present in his 1959 'Review of B. F. Skinner's Verbal Behavior', but sprouted in 'Quine's Empirical Assumptions' of 1968 (largely a response to a famous set of arguments in Quine's in 1960, *Word and Object*). Quine responded immediately, again in 1975, and defended his broadly behaviourist orientation in books of 1992 and 1996. Chomsky, for his part, devoted many pages of his 1975 book *Reflections on Language*, his 1980 book *Rules and Representations* and in the essays of his 2000 *New Horizons* to the dispute.

You may know Chomsky from YouTube. He appears in videos recorded from the late 1960s to the present day, speaking, lecturing and answering questions about politics and social issues, but also of course about linguistics and cognitive science, often delivering what at the time were the latest developments of his theory, and often at a relatively non-technical level (see the YouTube channel *Closer to Truth*, the four-part series 'Noam Chomsky's Reflections on Philosophy and Linguistics'). There is also the 1971 video of his appearance with Michel Foucault, which is not the communication failure you might suppose. In recent years he has spoken about ChatGPT, making the somewhat obvious point from his point of view, that mimicking or simulating an aspect of human intelligence tells us little about how humans do it. It is a point that goes way back in the history of computers. In 2000 Chomsky quotes Turing himself, in 1950, as saying the question of whether machines think 'may be too meaningless to deserve discussion' (Turing 1950, p. 442; quoted in Chomsky 2000, p. 44). The reason of course is that *thinking* is not a precise scientific concept, and has always been used under the ordinary presupposition that the beings in question are human, much as Wittgenstein – a friend of Turing – thought: 'We can only say of a human being, or what is like one that it thinks' (Wittgenstein 1953, §360). The famous Turing Test is a proposal for sharpening our concepts, not an hypothesis as to the real meaning of 'to think'.

● CHAPTER SUMMARY

The basic idea of Chomskian Universal Grammar or UG is to elucidate the deepest principles underlying the human language capacity. It is intended as not only

descriptive of human language but as explanatory of actual grammar, including not only our reflective judgements of grammaticality but reflective judgements concerning various types of grammatical relationship. The puzzles involved in carrying it forward are legion, complex and enormously difficult, as evidenced by the main changes – and the main advances – that Chomskian linguistics has undergone since it first appeared in the 1950s.

The so-called 'Poverty of the Stimulus' is a central empirical reason for thinking that UG is indeed real, despite what might seem at first blush to be the baffling variety of human languages. The POS-argument is simple, at least in its most basic form: in a very short time, on the basis of limited and degraded data, the child acquires a very complex language, making it extremely unlikely that the child does it by any other means than activating and calibrating a faculty that grows as a matter of genetic programming in the child's brain. Language capacity should be classed with the visual or auditory capacity, neither of which is acquired from without but grows within, requiring only minimal stimulus. UG is Chomsky's partial answer to 'Plato's Problem'. It marks Chomsky out as a thoroughgoing Naturalist, seeing the problems as ultimately part of Biology. He sees parallels with older forms of Rationalism, but with genetics substituted for theology.

All human languages exhibit compositionality – the derivability of phrases and sentences of arbitrary complexity and length from finite resources. This is a fact of UG, not just a fact of the particular languages as actually spoken – Swedish, Farsi, Japanese and their dialects. Chomsky introduces the concept of 'E-language' or 'external language' for these: E-languages are vaguely individuated, and often so only relative to the interests of the speaker (often the interests will be political, social or historical). Of scientific interest rather is an individual's particular competence with their language, their 'I-language', their 'internal language'.

An important 'linguistic universal' is the fact that human languages exhibit phrase structure, not just serial structure. The latter is a logically possible way for language to be, but in reality we understand sentences by looking out for the main verb (the main verb-phrase) of the sentence, wherever it is in the sequence of words. Similarly phrase categories such as noun-phrases hang together in particular ways, with the phrases hanging together in particular ways. To show these relations of structural dependence can be achieved either by bracket-notation or by tree-diagrams.

The base structure of grammar must be distinguished in some way from higher-level grammatical structure (equivalently: deep or D-structure must be distinguished from surface or S-structure). For a variety of reasons, the basis – called for a time 'phrase structure grammar' – may be adequate for mapping the grammatical sentences of a language, but is inadequate for capturing the other relations among phrases and sentences which are plainly matters of grammar, such as the relation of a sentence to other sentences which 'say the same thing' (such as passive vs. active), or to the corresponding types of questions. In later work, 'mono-dimensional' alternatives were proposed, but the important thing for our purposes is simply to appreciate the shape of the problems and the suitability of UG to cope with them.

In later work Chomsky replaced constellations of rules – sometimes seemingly specific to a language – with the notion of Principles and Parameters, with the former being fully universal and the latter comprising an array of two-way settings, which can be set with minimal exposure to a language. The success of such a programme would mean greater and simpler application to the various human grammars at the same time as greater explanatory value.

Chomsky's notion of 'meaning' is an internalist notion, not an externalist one. This means that the meanings of phrases – as products of LF (Logical Form) and the conceptual faculty – though rich and variegated, do not depend for their identity on how the creature happens to be causally embedded. On the side of LF – essentially the side of complex UG syntax – we know a great deal. On the cognitive side, Chomsky allows that the science is comparatively undeveloped, but contributes certain POS arguments to show that externalist accounts are unpromising.

It is commonplace to hold that the very *raison d'être* of language is communication, but Chomsky demurs. For if at a deep level human language is grounded in genetics, there is no necessity in language's being in fact used for communication. More to the point would be that its function is *thought*, but in fact it is unjustified to speak of 'the function' of a genetically determined trait. The spine for example has evolved playing many functions, sometimes in competition with each other. Indeed there is some evidence to suggest that the one trait that distinguishes human language from other systems of communication throughout the animal kingdom – the aforementioned trait of compositionality – was due to a single genetic mutation. Chomsky identifies the point at which UG goes compositional in the rule 'Merge', the principle for any phases X and Y, the unitary object is formed comprising X and Y, set-theoretically the pair {X, Y}.

UG is a theory of the relatively stable system of linguistic *competence*, not the spectacularly complex and seemingly highly variable domain of *performance*, a theory of which would tell what people x will say in situation y, for variables x and y. Connected with this is that since the facts of UG are ultimately neurological, UG is not bound by the facts of human behaviour – although it would be astonishing if behaviour were not a good way in.

● STUDY QUESTIONS

- 1 The notions of a Verb (V), Noun (N), Adjective (Adj) and Preposition (P) are probably well-known to you (prepositions include words like 'from', 'with' and 'in' etc.). A noun-phrase (NP) is a complex of words that play the role of a noun, like 'the red book'; likewise for VP, and PP (Adjectives can be treated as parts of VPs or NPs). Determiners (det) – 'the', 'an', 'all' etc. – always occur as part of a noun phrase. Assuming that the top two levels are S (sentence), and NP and VP, and perhaps using a children's book as a source of examples, construct a few (elementary phrase structure) tree-diagrams. Remember that a VP for transitive verbs has the structure VP(NP).

- PRIMARY READING

● SECONDARY READING

- NOTE

- 1 Strictly we should distinguish traces proper – positions in S-structures from which items have moved – from marked locations in D-structures into which a noun phrase must be moved to generate the S-structure. For example, the S-structure ‘The egg was boiled *t*’ is based on the D-structure ‘*e* was boiled the egg’. This comports with the principle discussed below (under θ -theory) that the lexicon includes the entry ‘boil [___ NP]<agent, patient>’, covering all forms of the verb ‘to boil’. The transitive verb demands a subject, even if it takes the form of an empty category.

13

• modern directions I

In this chapter, we will consider four areas on which contemporary philosophers of language have conspicuously made their mark (but we will often set the stage with the sorts of concept and theory we've already examined in this book): (1) assertion; (2) context-relativity; (3) fictional entities; and (4) inferentialism. We mean merely to whet the reader's appetite for these topics; by no means will our treatment be exhaustive. Chapter 14 will consider three more.

● ASSERTION

Frege characterised *judgement* as 'acceptance of a proposition as true'. He characterised **assertion** as the 'verbal manifestation of judgement'. He could have seized on other verbs which perhaps express slightly different concepts – *to say, to claim, to state, to affirm, to declare, to proclaim, to announce, to put forward, to avow* and so on. But most philosophers have at least implicitly taken Frege's lead in thinking that assertion is the central concept here.

In Chapter 9, we followed Austin in maintaining that the speech-act of assertion is governed by conventions. For reasons of space we will leave aside the idea that assertion is explicable non-conventionally in terms of the psychology or intentions of the asserter – the views deriving from Grice's seminal papers 'Meaning', of 1957, and 'Utterer's Meaning and Intentions', of 1969, which among other things aimed to define assertion in terms of the communicative intentions of the speaker (see the 'Study questions' at the end of the chapter for a taste). Instead, we will assume that the speech-act of assertion exists at least partly as a conventional kind of act, a socially constituted practice. Michael Dummett puts it well:

the linguistic acts should be classified as conventional actions, not as the external expression of interior states. Assertion, for example, is to be explained

in terms of the conventions governing the use of those sentences which are understood as having assertoric force, not as the utterance of a sentence with the intention of expressing one's interior act of judgement (or interior state of belief) that it is true.

(Dummett 1973, p. 311)

So our question: what exactly are the conventions that characterise this particular speech-act?

We first must make a certain distinction, following Searle (who followed Elizabeth Anscombe and John Rawls), between *regulative* rules and *constitutive* rules. Regulative rules are rules which govern a pre-existing activity, such as 'Drive on the left', a rule that governs driving in the UK. Widespread violation of this rule would not threaten the very existence of driving in the UK (even if it would threaten the existence of the drivers). Constitutive rules, by contrast, are ones whose existence is necessary for the existence of the practice or act in question. The rules that define a game such as chess or football (soccer) make it possible to *castle* or *score a goal* rather than just rearrange two pieces on a board or kick a ball into the net. More generally, if there were no rules of chess or football, chess or football, in the sense we know them, would not exist. This is what it means to say that a form of life is conventional. It is the existence of constitutive rules that make various social institutions or forms of life possible, such as *buying and selling*, *ownership*, *marriage* and so on.

The distinction has been challenged, but here we'll assume that the distinction is well taken. The idea is that assertion is bound by certain constitutive rules. Consider now an instance of what is known as **Moore's paradox** (named after G. E. Moore, an influential philosopher who was roughly contemporaneous with Russell). Suppose Gia says:

There is cheese in the refrigerator, but I don't believe it.

There is something wrong with Gia's utterance. But notice that the *sentence* uttered is not logically contradictory (as in 'p and not-p'), for obviously it is a possible state of affairs that there is cheese in the relevant refrigerator, when Gia does not believe that there is cheese in the relevant refrigerator. What is wrong is that she, in some sense, 'represented herself' in speaking as she did, as accepting, and therefore as believing, the content of 'There is cheese in the refrigerator' – at the same time she explicitly said that she does *not* believe that very content. The infelicity is not at the semantic level. The infelicity seems to be at the pragmatic level, appertaining to the utterance: it is part of the very nature of assertion that one should say only what one believes. Thus it appears that the following is a constitutive rule of assertion, such that to violate it is to violate a convention governing assertion:

(Belief-rule) One must assert that p only if one believes that p.

But notice further that if Gia simply asserts 'There is cheese in the refrigerator', and she does not believe that there is cheese in the refrigerator – perhaps she in

fact knows that there is no cheese in the refrigerator, in which case she lies or tells a fib – then she violates such a constitutive rule but makes such an assertion all the same. The belief-rule is automatically invoked wherever one makes an assertion, but one can disobey such a rule or norm, just as promises can be broken. In promising, one necessarily undertakes an obligation to fulfil it – it is constitutive of the act of promising that one incurs an obligation to keep it – and likewise it is constitutive of the act of assertion that one should believe the content that one is asserting. A violation of the belief-rule is an act of lying, untrustworthiness or insincerity (storytelling, we assume, is not an act of assertion, at least not straightforwardly so).

So far, the observations might seem to steer us to Gricean territory. Is not a violation of the belief-rule a violation of the cooperative principle, in particular the maxim of quality? Yes. But Gricean maxims are merely regulative, or rules-of-thumb. We are now discussing possible *constitutive* rules, rules which, as we say, *define* what it is to make assertions.

Still, a look at Grice's maxim of quality might suggest other possibilities (the maxim is: 'Do not say what you believe to be false; do not say anything for which you lack adequate evidence'). Thus:

(Truth-rule) One must assert that *p* only if *p*.

(Justification-rule) One must assert that *p* only if one has a justified belief that *p*.

Notice that, like the belief-rule, these do not specify sufficient conditions for when you ought positively to assert something. They specify only necessary conditions of making felicitous assertions; they represent constraints on assertion that one is expected to obey. Combining them, we have:

(JTB rule) One must assert that *p* only if one has a justified true belief that *p*.

Justified-true-belief (JTB) used to be thought of as *knowledge*. But Edmund Gettier famously pointed to cases of JTB which are *not* knowledge. For example, I might believe that there is a dog on the other side of the fence. Although I can't see it, I justifiably believe it because I can plainly hear a dog barking in the right direction. In fact, unknown to me, what I'm hearing is not a dog but a recording of a dog. There is, however, quite unbeknownst to me, a dog lounging on the other side of the fence, who keeps silent. So it seems that I have a justified, true belief that there is a dog on the other side of the fence, but I do not *know* there is a dog on the other side of the fence. That I'm in this peculiar situation seems quite accidental: the justification for my belief (the recording) has nothing to do with what makes the belief true (the lounging dog).

At first it seems a definition of knowledge just needs a little tweaking to cope with such a case, but the history of attempts at patching up the definition show the intractability of the problem, as the variations to Gettier's initial problem cases multiply. It has led Timothy Williamson to argue that knowledge *cannot* be analysed in this way, to reckon it *primitive* – that is, as *indefinable* – and to propose the following constraint on assertion:

(Knowledge-rule) One must assert that *p* only if one knows that *p*.

It is a constitutive rule. And, again, there are plenty of assertions where the knowledge requirement fails. Nevertheless, in asserting something, an asserter is subject to the rule in exactly the same sense that a promiser is subject to the norm that they ought to keep the promise.

The proposal predicts that a speech-act involving the following is infelicitous (the example extends Moore's paradox):

There is cheese in the refrigerator, but I don't know it.

Williamson's proposal has won many adherents but also has had its critics. On the positive side, the knowledge-rule would seem to dovetail with the thought that assertion is a way of *transmitting knowledge*, and with the thought that if you assert something you're thereby committed to *defending* it. But lesser norms – the truth-rule, or even just the belief-rule – have been claimed to do this work. And it has also been claimed that we go wrong at the start in supposing that assertion is the central concept of the philosophy of language or pragmatics; the declarative speech-acts we actually undertake are nothing like as uniform as that. Telling, reminding, warning, prompting, advising, counselling, expressing, directing, describing, reporting and so on – the differences are often subtle, but perhaps each brings with it a distinctive set of norms.

● CONTEXT-RELATIVITY

You're familiar by now with indexicals. They include 'I/me', 'you', 'we/us', 'he/him', 'she/her', 'now', 'then', 'today', 'yesterday', 'here', 'there', 'that', 'this' and so on; also there are mixed indexicals such as 'your mother's cat'. Such expressions are context-sensitive or context-relative: only with respect to particular contexts do they have referents, and they refer to different objects depending on the context in which they are used (which includes time, place, identity of speaker and intended audience, identity of any demonstrated objects).

We mentioned very early on, only to set them aside, words like 'big', 'rich' and 'near', which seem sensitive to context as well.¹ For example:

(1) That's a big one!

said of a rat, seems to assign a different range of size to the relevant rat from the one that the same sentence said of a hippopotamus assigns to the relevant hippo. Indeed, the first might be true and the second false, if the size of the intended hippo were distinctly less than that of the average hippo – even were the intended hippo to dwarf the big rat. It can be small 'for a hippo'. It seems that the standard for bigness, or the relevant *comparison class*, shifts, depending on context.

We seem to see analogous shifts in standards for 'rich'. We say that a person is not rich if the person is a lower-middle-class San Franciscan (where the median

income in 2021 was US\$77,267), but if the person earned the same income and so on but was an Ethiopian living in Ethiopia (where the median income in 2023 was US\$1,764), we'd not demur from the proposition that the person is rich. Similarly, looking at a globe, considering potential cities to visit, an utterance of 'That's near' seems to invoke a much less demanding standard of proximity than using the same sentence when talking of potential restaurants to walk to for dinner.

The key in such cases – comparative adjectives and prepositions – seems to be the *implicit relativity* to a standard or comparison class. A likely hypothesis is that words such as 'big' carry with them a tacit indexical that refers to a standard or comparison class. Made explicit, we can write *big_x*. (Indexicals are often thought to be just free variables with their domain restricted by a certain rule.)

Another sort of case of apparent context-sensitivity is sentences containing quantifiers or quantifier-phrases such as 'everyone' (we've mentioned this before, in the first section of Chapter 8). Consider:

(2) Everyone was rude.

Suppose Sally says this when emerging from a party. It's plausible that this utterance of (2) is true if and only if every person *at the party* was rude (or everyone at the party who interacted with Sally was rude in that interaction), not that every living person anywhere and anytime was rude. It seems that the context is set partly by the intentions of the speaker. We can imagine other contexts of utterance of (2) being different just over the intended scope of the generalisation; the quantifier 'everyone' might range over, say, *the people Sally interacted with during her stay in New York*.

It stands to reason that just as we should think of comparative adjectives as carrying with them a tacit indexical that refers to a comparison class, we should think of quantifiers as carrying with them a 'domain restrictor', in effect a tacit indexical referring to a class of objects. Indeed, this seems inevitable when we look back at the use of definite descriptions, understood in Russell's way. 'The F' is in effect a quantifier: 'The F is G' is true just in case *there is* exactly one F, and every F is G. If there is exactly one F, then the definite description denotes a certain object, as Russell says, or, as we can say more sloppily, refers to a certain object. But, for example, if the domain of 'the cat' were taken to be all living cats whatsoever, then it would never succeed in securing a referent (except in a sad little world with exactly one surviving cat). But, of course, it does typically have a referent; sometimes, not in the sad little world but in the actual world, we can assert 'The cat is on the bed' with perfect veridicality. So we attach a subscript to 'the cat' – and to 'the bed' while we're at it – representing the envisaged utterance as one of 'The cat_x is on the bed_y', where the domain restrictors for x and y are *is owned by the speaker* and *in this flat*, or something like that. Formally, in Russell's way treating 'is on the bed' simply as a predicate 'B_': $\exists_x y [Cy \ \& \ \forall_x z (Cz \rightarrow z=y) \ \& \ By]$.

Yet another case of implicit context-relativity is a familiar one. The time is not rare that 'It's raining' is true in Edinburgh but false in Athens. The truth-value given a time of utterance can vary with the place of utterance and also, given the place of utterance, can vary with the time of utterance. (One may also speak of the weather in Athens from Edinburgh, by explicitly saying 'In Athens, it is raining'.) Similarly, it varies with 'It's windy', 'It's noisy', 'The air smells of jasmine', 'It's bright' and so on.

Philosophical concepts I: value. Also familiar, and the source of some contentiousness, is the apparent implicit relativity of at least some value statements. Joe says 'The Beatles were the greatest!' Karen answers 'The Rolling Stones were the greatest!' Do they disagree? Are the two positions mutually contradictory? They might well argue over it. But maybe, in this case, they ought to accept that their respective statements contain an implicit indexical, so that Joe says that the Beatles were greatest *for him*, and Karen says that the Rolling Stones were the greatest *for her*. If so, both positions can be true.

The same trick can be pulled with respect to any statement of value – of ethical value, of aesthetic value, of matters of taste. Perhaps the trick is justified in the case of Joe versus Karen. More decisively, it is justified in many cases of gustatory taste, like whether anchovies taste good: they are good *to some* and not good *to others*, and it seems that only an over-the-top objectivist or food snob would say that there must be a fact of the matter concerning who is right. Some like them, others don't, end of story. It seems far less justified in cases of the most serious ethical concern, such as whether slavery is wrong. No one can get out of a professed fondness for slavery by saying 'But it isn't wrong *for me*!' It is the middle ground that is contentious, where battles are worth fighting over whether the apparent disputes are real, or whether they ought to be deflated by inserting an indexical, making *n*-place predicates into *n+1*-place predicates.

It is not only the speaker who can be the relevant relatum in the relativised form of statements of value. Very frequently we want to say that *x*, in the form 'A is valuable to *x*', is not the individual speaker but a class of persons, such as the members of a culture, a nation or other social group.

Philosophical concepts II: knowledge. Imagine saying:

(3) I know she's in the bathroom.

(You just saw her go in, and the door has remained closed.) You have also been known to say, in a philosophy class:

(4) I never know that it isn't all a dream.

Could both be true? On the one hand, you might think that if (4) is true, then (3) can't be true. If you never know that you're not dreaming, then in particular you don't know that your experience of your friend entering the bathroom was not dreamed; in which case, it seems, you don't know that she is in the bathroom.

On the other hand, maybe the standards for knowledge shift depending on the context, in such a way that in the context of a classroom discussion of sceptical scenarios and such like, one rightly assumes a very strict standard for knowledge, but one rightly relaxes the standard in ordinary circumstances such as that of saying (3). In that case, both statements could be true.

Thus according to *epistemic contextualism*, the exact proposition expressed by a claim of the form 'S knows that p' varies not just according to who S is and the content of p but upon the standards in play. Once again, we could suppose that 'knowledge' sports a hidden indexical, a variable fixed by the context of utterance: *know_x*, where the parameter 'x' is replaced in context by a phrase describing the standards or kind of knowledge that prevail in that context or which are intended in that context. So one cannot know *simpliciter* that she's in the bathroom, but one can *know for ordinary purposes* that she's in the bathroom, but not *know_{strictly}* that it isn't all a dream. Perhaps one *knows for ordinary purposes* that one is not dreaming, but doesn't *know_{strictly}* that one is not dreaming.

A more extreme view – generally known as epistemic relativism rather than epistemic contextualism – involves the claim that there is no such thing as *the* truth-value of the proposition expressed by an utterance such as 'I know the floor is flat', even with all facts fixed pertaining to its context of utterance. It depends also on the perspective and purposes of the one ascribing the truth-value, who may be quite outside that context. Assuming that the utterance was of a particular floor, the self-same proposition can be true relative to ordinary purposes but false relative to the exacting standards required by carrying out an experiment in physics. In general, the epistemic relativist holds that even in cases where 'A knows that s', with A, s and the context of utterance given – thus it expresses a given proposition or content – still the truth-value can shift depending on the standpoint from which the proposition is evaluated. The position is not one of *subjectivism*: once a particular set of standards is chosen, one can still be wrong in estimating a statement's truth-value (and the question of which standards are rightly invoked is often, if not always, rationally constrained by features of the context of the assessor of the knowledge-claim).

Philosophical concepts III: counterfactual conditionals. In Chapter 7, we briefly considered **counterfactual conditionals**. They are not amenable to being explained in terms of simple truth-tables. The truth-values of p and q do tell us, at least arguably, the truth-value of the ordinary ('indicative') conditional 'if p then q'. The truth-table tells us that (i) if p is *true*, then 'if p then q' is true if q is true, and false if q is false; (ii) if p is *false*, then the conditional is true no matter what the truth-value of q. Thus the conditional is equivalent to 'Not p, or q', which is true so long as p is false or q is true.

There is room to dispute the truth-functional account of indicative conditionals, but such an account definitely will not work in the case of the counterfactual or subjunctive form 'If p *were* so, then q would be so' (or 'If p *had been* so, then q would have been so'). In the case of a typical utterance of a counterfactual conditional, we know that p is not the case (and, typically, that q is not the case), but that doesn't tell us whether or not the conditional is true.

For example:

- (5) If Roald Amundsen had not been the first to reach the south pole, then Robert Scott would have been the first to reach the south pole.

Amundsen *was* the first, and Scott was not, but that does not settle the truth-value of (5).

According to David Lewis' influential account of counterfactual conditionals, the truth-value of such a conditional is determined as follows. Consider the possible world most like the actual world except for the antecedent's being true (where Amundsen was not the first to make it to the south pole); in that world, with *everything possible the same as in the actual world*, is the consequent true? If yes, the conditional is true; if not, it's false.

This appears not to help with examples such as this (from Quine; it is the Korean War, of 1950–53, that is the subject of discussion):

- (6) If Caesar had been in command, he would have used the atom bomb.
(7) If Caesar had been in command, he would have used catapults.

At first blush, it seems that these can't both be true. But according to Lewis, each of these conditionals can be true *relative to the context of utterance*. In particular, if we keep fixed Caesar's general ruthless character but grant him twentieth-century weapons of war rather than restricting him to those available in the first century BC, we'd select the corresponding possible world and presumably we'd find (6) to be true but not (7). But if we wished to confine him to first-century weapons of war while retaining his character, then it's likely we'd choose (7) and discount (6) as false. Thus the truth-conditions of counterfactual conditionals depend partly on the context of utterance – in particular on the subject, tacit assumptions and interests of the conversation.

Radical contextualism. We have identified many cases of context-dependence that do not involve explicit indexicals (see note 1 for more). You may well have wondered: when does it stop? Maybe *all* expressions are sensitive to context! Our friend Austin seems to have believed this. He famously puts the point as follows:

if you just take a bunch of sentences ... impeccably formulated in some language or other, there can be no question of sorting them out into those that are true and those that are false; for ... the question of truth and falsehood does not turn only on what a sentence is, nor yet on what it means, but on, speaking very broadly, the circumstances in which it is uttered. Sentences as such are not either true or false.

(Austin 1962, 110–111)

And gives this example:

Consider ... 'Lord Raglan won the battle of Alma', remembering that Alma was a soldier's battle if ever there was one and that Lord Raglan's orders were

never transmitted to some of his subordinates. Did Lord Raglan then win the battle of Alma or did he not? Of course, in some contexts, perhaps in a school book, it is perfectly justifiable to say so – it is something of an exaggeration, maybe, and there would be no question of giving Raglan a medal for it ... ‘Lord Raglan won the battle of Alma’ is exaggerated and suitable to some contexts and not to others; it would be pointless to insist on its truth or falsehood.

(143–144)

Joining him, at least in the spirit of the idea, are the later Wittgenstein, Stanley Cavell and John Searle, and recent philosophers of language such as Charles Travis, Cora Diamond and Avner Baz. Take a simple, innocuous example, so that it will become at least plausible that no sentence is quite context-free in its truth-conditions (the example is adapted from a discussion by Travis). I say to you, of your very expensive new shoes, ‘Your shoes are under your bed’. In fact, they are not even in your room; in most contexts of utterance, what I say would be false. But suppose the shoes are located one floor down – in a region between your bed and the centre of the earth; yours is a waterbed that has sprung a considerable leak, and I said what I said in order to alert you to the danger posed to anything below the bed, including your expensive new shoes. The idea is that I would have spoken truly.

● FICTIONAL OBJECTS

In Fregean semantics, we had the difficulty that if a name does not denote anything, then, although sentences containing it can express thoughts (they can express complete senses), those sentences fail to have truth-values (except when they occur inside the scope of ‘believes that’ etc.). Calamitously, an evidently true sentence such as ‘Odysseus did not exist’ comes out as not true or false, assuming that the story of *The Odyssey* is fanciful and the figures who appear in it are fictional (which we will do in this discussion). Russell’s theory did better on this, so long as we suppose that ordinary proper names are definite descriptions in disguise. The definite description that abbreviates ‘Odysseus’, on a particular occasion of utterance, would be something like ‘The Greek warrior who blinded the Cyclops’, where the description is analysed in the way that Russell proposes (see Chapter 3, ‘Applying the theory of descriptions’). Then ‘Odysseus’ wife waited for him for twenty years’ would come out false, as would ‘Odysseus’ wife did not wait for him for twenty years’ (on a wide-scope reading of the description substituted for ‘Odysseus’).

It is perhaps artificial to think names (in a given context of use) are ‘really’ descriptions. But what’s worse, it seems that of this pair of sentences –

- (8) Odysseus’ wife waited for him for twenty years.
- (9) Odysseus’ wife gave up waiting for him after three months and married the most handsome of her suitors.

– both are false, on Russell’s reckoning. But surely there is something right about (8) and wrong about (9). Indeed, many people, certainly those uninfected by philosophy

of language, wish to assert (8) but deny (9); they would say that (8) is *true* whereas (9) is false.

Possible objects. A natural move would be to say that although Odysseus is not an actual person, he is a *possible* person. There are possible worlds in which Homer's tale is real history, in which Odysseus exists and did the things recounted by Homer, such as being held by the beguiling Calypso, blinding the one-eyed monster Cyclops and returning to his faithful and resourceful wife Penelope after twenty years of war and adventure. So long as we imagine a tacit 'In possible worlds at which Odysseus exists ...' tacked on the beginning, this move would seem to make (8) true and (9) false, just as the doctor ordered. The proposal then is that fictional entities are just a subset of possible entities, most of which exist only at non-actual possible worlds (and most of which will never be thought of, unlike Odysseus).

One problem with the idea as it stands is that it's under-described. Take a proposition *p* involving Odysseus which is consistent with Homer's story, but whose falsity is also consistent with Homer's story (suppose that it's quite incidental to the story, such as *Odysseus ate seventeen olives the morning when he blinded the Cyclops*). Call those worlds at which both Homer's story is true and *p* is true W_p , and those worlds at which Homer's story is true and *p* is false $W_{\text{not-}p}$. Which is the real Odysseus? The Odysseus-type-being of W_p or the Odysseus-type-being of $W_{\text{not-}p}$? For real existing people there is a fact of the matter about such things as whether on a given morning the person ate seventeen olives; but in the case of Odysseus, no such fact has been determined.

A likely counter is to suppose that when we talk of Odysseus, we are really talking of a *class* of possible worlds, namely those in which Odysseus possesses the attributes explicitly ascribed to him by Homer (and all those attributes that follow logically from the set of attributes ascribed to him by Homer). However, even the propositions expressed by Homer in *The Odyssey* concerning him are not sacrosanct. We can easily imagine Odysseus as being exactly as Homer describes him except that he was tied to the bow of his ship rather than to the mast of his ship when he heard the Sirens singing. It does not seem as if in doing so we are no longer thinking of Odysseus. But, then, how many of the things that Homer relates of Odysseus can be imagined otherwise, without ceasing to be stories which feature Odysseus? This is a classic 'How many angels can dance on the head of a pin?'-type question; a theory which suggests that there have to be answers to palpably bootless questions can presumably be set aside.

Neo-Meinongianism. This is a modernised version of the scheme put forward by Alexius Meinong in the late nineteenth century. We can get partial relief from the above problem if we simply accept that fictional entities *are* indeterminate in ways that ordinary objects are not and thus refuse to identify them in any sense with definite possible objects. Odysseus, on this view, did not exist. He did kill the Cyclops and so on *but neither ate seventeen olives on the day of killing the Cyclops nor did not do so*. Unlike possible objects, he is indeterminate in this respect. The view splits into two variants. The first variant holds that, like ordinary objects, fictional objects are *concrete* (but still nonexistent). This fits with the idea that since Odysseus blinded the Cyclops – a physical action – he must have been concrete. The second variant

holds that fictional objects are not concrete objects but abstract objects of a certain kind. Think of *generic* individuals, individuated by certain combinations of properties: *the quarterback of the San Francisco 49ers*, for example, is a combination of properties that is realised by various concrete individuals over time. Rather than thinking in terms of definite descriptions, as Russell recommends, this can be taken as an abstract object that remains the same through time (so the relation of, say, Joe Montana to the role, is one of *fulfilment* for a certain period) – similarly with Odysseus. Abstract or concrete, the main thing according to neo-Meinongeanism is that nonexistent objects are often quite spare in their properties; just a few properties, even perhaps just *one* property, suffices for the being of such objects.

Related is the strategy of appealing to *impossible worlds*, with impossible objects amongst the inhabitants of the worlds. Yes, they are a thing! With some standing and interest among logicians and metaphysicians; I refer you to Berto and Jago, 2019.

Creationism. Thinking of fictional entities in these ways seems not fully to comport with such thoughts as that Homer's *The Odyssey* was composed about 2800 years ago and therefore that Odysseus is a *contingent* object about 2800 years old. Remember that according to Kripke, the referent of a typical proper name is determined by a chain of communication, or referential intentions, going back in time to an original act of baptising an object, paradigmatically an act of standing over the object and saying something of the form, 'Let this object be called "N"'. Kripke suggests that *fictional* names can be thought of as being introduced analogously. Take the more recent case of Jane Austen's Mr Darcy, of her novel *Pride and Prejudice*. When Austen wrote down this story, she thereby instituted a practice of speaking of a certain Mr Darcy, who did various things along the way to capturing the heart of one Elizabeth Bennet. Her creation of Mr Darcy is the creation of an abstract object, but, much like the equator, this object is a contingent object, has a beginning in time and furthermore is partly a linguistic, conceptual or intentional entity. And it is necessarily *her creation*. Therefore it has some ammunition to ward off the above objections afflicting the possible objects theory: a description is only of Mr Darcy if it is true, and causally connected in the right way, of *the creation of Jane Austen*. So there is no puzzle of nonexistence: Mr Darcy – like Snoopy, like Odysseus – does exist but in the form of various descriptions, intentions, images and roles played by actors in films, not as flesh and blood. The evident tension perhaps is that it appears that is true that Darcy is only imaginary yet that he genuinely loved Elizabeth.

Fictional operators. Another idea is to suppose that when one speaks of Odysseus, there is an implicit or tacit operator that governs the speech-act, which might spelled out as 'According to a story [...]', or 'According to the story commonly attributed to Homer [...]'. The operator might be taken as a hyper-intensional operator, in which case the truth of such speech-acts is grounded in the senses or conceptual content of the sentences (including implicatures and some of the presuppositions) within the scope of the operator. According to an interesting wrinkle on this approach known as pretense theory, we can understand the operator in terms of the psychic attitude of *pretending*: the operator instructs us to adopt that attitude. This fits with the apparent fact that children have no trouble getting that stories aren't real.

All of these views struggle over certain true, mixed statements like ‘Snoopy often lay on top of his dog house; the image of him doing so is well known’ or ‘Plato knew of Odysseus but naturally did not meet him’. Where the fictional operator is held to be in play, these statements seem to call for the operator to lapse, sometimes within a single sentence, in a seemingly *ad hoc* manner.

● INFERENCEALISM

The classical theorist of meaning regards pragmatics – the use of language – as built upon semantics. Semantics fundamentally involves reference: singular terms paradigmatically *refer* to objects; predicates *refer* to properties and relations. Once one learns the semantics, one can employ the words in various kinds of speech-act, including the drawing of inferences.

An alternative, articulated by Willard Sellars (1912–89) and Michael Dummett (1925–2011), and championed recently by Robert Brandom and Jaroslav Peregrin, by contrast regards pragmatics, and in particular the drawing of inferences, as fundamental. It’s called inferentialism. Consider the logical particle ‘and’ (where its role is that of connecting whole sentences, not its role in forming collective subjects, as in ‘Laurel and Hardy’). Table 13.1 is a truth-table for the word, letting ‘P’ and ‘Q’ stand for any declarative sentences.

Table 13.1 Truth-table for ‘and’

P	Q	P and Q
T	T	T
T	F	F
F	T	F
F	F	F

This shows that if P and Q are both true, then ‘P and Q’ is likewise true; otherwise ‘P and Q’ is false. The standard inference rules for ‘and’ – employed in reasoning, or formally in proofs or derivations – are as in Table 13.2.

Table 13.2 Inference-rules for ‘and’

P, Q	P and Q
P and Q	P, Q

These tell us that anytime one has P and also has Q, one may infer ‘P and Q’; and if one has ‘P and Q’, one may infer P and likewise may infer Q.

The classical theorist regards the information in the truth-tables as fundamental; the inference-rules – instructions for how to *use* the word – are *based* on the truth-tables, which specify the *meaning*, the semantics, of ‘and’. But for the inferentialist, it is the other way around. It is inference-rules that are fundamental: they set forth what one has to know, if only implicitly, to be credited with mastery of the word ‘and’. If one possesses such mastery, nothing else is required for grasping the meaning of ‘and’; one understands the word just insofar as one performs with the word as specified by the inference rules. The rules make explicit what was originally a matter of practice, of the use of words – a matter of *knowledge-how* as opposed to *knowledge-that*.

Such are ‘inter-linguistic’ rules; ‘intra-linguistic’ rules comprise ‘language-entry’ and ‘language-exit’ rules, which come to the fore with non-logical sorts of word. Instead of explaining the word ‘red’, for example, in terms of its *reference* to a certain property or the *truth* of ‘x is red’ for certain objects, we specify rules such as: if an object is seen to be red, one may assert ‘It is red’; if commanded to ‘Get the red one’, one should get the red one (in this case there are also such inter-linguistic rules as: from ‘It is red’ one may infer ‘It is coloured’, ‘It is not blue’ and so on).

This approach disenchants the concepts *reference* and *truth* in favour of those fundamental to the statement of such rules as just described; in particular the general form of such rules tells one what one is *obligated* to do, and what one *may* do, with words; such are **normative** concepts, not semantic concepts. Indeed, the words ‘refers to’ and ‘is true’ can themselves be subjected to the same treatment. Very roughly, one is entitled to assert *P is true* just in case one is entitled to assert *P*; and similarly one is entitled to assert *the referent of ‘b’ = A* just in case one is entitled to assert *b = A*.

Only some inferences one is prepared to draw can plausibly be explained as justified by one’s mastery of words. For example, if one happens to know that one has a sandwich in one’s backpack, one’s inferring ‘Here is some lunch’ from ‘Here is my backpack’ is justified, but it is not justified merely by one’s mastery of words. But some inferences can be seen in that way – for example, the inference from ‘A is north of B’ and ‘B is north of C’ to ‘A is north of C’. The idea is that the sum total of inferences of this sort constitute one’s linguistic mastery, one’s knowledge of meaning.

● CHAPTER SUMMARY

We assumed that assertion is rule-governed practice: that it is bound by a set of *constitutive* rules as opposed to *regulative* rules. Constitutive rules are ones such that without them, the practice could not exist; the rules of chess, for example, bring it about that certain moves of the pieces count as castling. Regulative rules are rules such as speed limits: the practice of driving would not cease to exist if the rule were cancelled. An initial thought is that the rule for assertion is that one should assert *p* only if one believes that *p*. An asserter is necessarily subject to this rule. Thinking of assertion in this way seems to resolve Moore’s paradox, that assertions of the form ‘*p* but I don’t believe *p*’ are not logically contradictory, yet an asserter of such a sentence is plainly in a logical pickle of some sort. The infelicity is analogous to the case

of one who says 'I promise to do A' but intends not to do A; one who asserts *p* or promises to do A is thereby subject to the norms that govern assertion or promising that one should believe *p* or do A. And just as there are insincere promises, there are lies (or attempted lies, where the person unwittingly says something true). Besides the belief-rule, other candidates include the truth-rule, the justification-rule and the knowledge-rule.

We examined whether context-dependence should be extended to other parts of language besides explicit indexicals. It seems that it should be, in the following cases. (1) Cases of implicit relativity such as 'big'; in such cases, the relevant index is a standard or comparison class. (2) Ordinary quantifier-phrases such as a teacher asking her class 'Everyone's ready?': the scope of the quantifier is implicitly restricted to the class. (3) Typical uses of 'It's raining' are implicitly restricted by an indexical 'here'. (4) Statements of value are often implicitly relative, as in 'It is good to me' rather than 'It's good'; the implicit relatum can sometimes be not the speaker but a class of people. We examined the idea that knowledge-claims are sensitive to context, distinguishing 'epistemic contextualism' from 'epistemic relativism'. We examined also Lewis's solution to a puzzle about counterfactual conditionals; such conditionals – well-formed conditionals with the same antecedents but conflicting though equally plausible consequents – in Lewis' view, are relative to an intended context which chooses between the two. We finally considered the proposal, often attributed to Austin, that all ordinary sentences, or virtually all of them, are context-sensitive.

Suppose we reject both Frege's and Russell's account of fictional names, and we want to take them as genuine, referring names. Thus we have to find suitable objects to which they can refer. One possibility is that they refer to merely possible objects. This has the advantage of not costing us much, since we seemed to need possibilities or possible worlds for other purposes. However, each possible object is fully determinate in a way that fictional objects are not; for example, a candidate for being identical with Snoopy was being born of a certain litter of a certain number of puppies, but nothing in the work of Charles Schulz tells us how many littermates Snoopy had. So perhaps we have to take the object as indifferently of that class of objects of which Schulz's stories are true. But even that is too strict, since we speak intelligibly of Snoopy as being Linus' dog, for example, when everyone knows that Snoopy is Charlie Brown's dog. Another strategy is to accept that fictional objects just are peculiar in that there is no fact of the matter concerning whether they have certain properties – i.e. there is no fact of the matter as to the number of Snoopy's littermates. According to Neo-Meinongianism, fictional objects do not exist, are not to be explained in terms of possible objects and have this admittedly odd feature of ontological indeterminacy. Another way is to take seriously and literally the plain fact that Schulz created Snoopy; Snoopy exists in the form of cartoon drawings, descriptions and television programmes; he is a contingent object, a cultural product.

Semantics, and in particular the referential power of linguistic expressions, is easily thought to be in some sense prior to the use of those expressions and in particular to their use in drawing inferences. Inferentialism has it the other way around, taking inference as prior with respect to reference. Indeed, to appreciate that certain inferential patterns are correct and other patterns are incorrect is just what it is to know

the meaning of the logical particle ‘and’ – likewise for the other particles of logic. If the notion of inference is widened to include the ‘inference’ of saying ‘That’s red’ if asked the colour of a red ball, and the taking of the red ball if asked to do so, the inferentialist can make a plausible case that the whole of language can be explained in this way, short-circuiting the apparent need for a concept of reference at all (but they can tell a similar sort of story as to our competence with words like ‘refers’). The crucial concept for the inferentialist is the general concept of a *rule* or *norm*. Language is not fundamentally for representing objects or states of affairs but is a system of rules; indeed, the former can in principle be explained in terms of the latter.

● STUDY QUESTIONS

- 1 Is Williamson’s knowledge-rule on assertion too strict? Think of examples in different walks of life (the classroom, the pub, heart-to-heart talks ...).
- 2 Are there apparent cases of *moral* disagreement that are really cases of semantical disagreement? What difference does it make?
- 3 Grice thought of assertion as a form of the wider category of ‘non-natural meaning’: S non-naturally means that p by the action A if and only if, for a certain audience H: (1) S intends A to produce in H the belief that p; (2) S intends H to recognise (1); and (3) S intends H to believe that p on the basis of (1). More roughly: to non-naturally mean something is to perform an action in such a way that your audience recognises that you want them to believe something on the basis of your having acted in that way. (Grice’s famous example: one draws a picture of Mr Y displaying undue familiarity to Mrs X and shows it to Mr X, thereby non-naturally meaning that Mr Y had been unduly familiar with Mrs X). Are there counterexamples to this view?
- 4 Le Verrier believed in the planet Vulcan (the small nonexistent planet whose orbit is inside the orbit of Mercury). Is Vulcan a *fictional* object? Try out each of the various accounts of nonexistent objects we discussed with respect to this case.
5. Another idea for nonexistent or fictional objects is that they really are mere ideas-in-one’s-mind. Possible objections: (1) different people talk of the *same* Santa Claus, which would be impossible according to the present suggestion; (2) Le Verrier would not accept that Vulcan exists *because he has certain thoughts*; he would say that Vulcan exists because of the astronomical facts. Are these objections decisive?
- 6 Is it really plausible that almost every sentence of English is context-sensitive?

● PRIMARY READING

The distinction between primary and secondary reading for these topics is drawn historically; indeed, the secondary reading generally will be more relevant. But still there are some ‘classical’ references that we can usefully distinguish as primary reading. Sources are listed by topic then chronologically (at the time when the piece was written) within each topic.

For assertion: Charles Sanders Pierce, 'Belief and Judgment' [1877]; Gottlob Frege, 'Thought' [1918], in *The Frege Reader* (1997), pp. 325–345; H. Paul Grice, 'Meaning' [1957], pp. 213–223 and 'Utterer's Meaning and Intentions' [1987], pp. 86–116, both in his (1989) *Studies in the Way of Words*; Michael Dummett, 'What Is a Theory of Meaning? (II)' [1975], in his *The Seas of Language* (1993b), pp. 34–93; John Searle, *Speech Acts: An Essay in the Philosophy of Language* (1969); Donald Davidson, 'Communication and Convention', in his *Inquiries into Truth and Interpretation* (1984), pp. 265–280 – also in his (2006) *The Essential Davidson*.

For context-relativity: John L. Austin, *How to Do Things with Words*, second edition (1962); Ludwig Wittgenstein, *Philosophical Investigations* [1953], fourth edition.

For fictional objects: Gottlob Frege 'On Sinn and Bedeutung' ['On Sense and Reference' 1892], in *The Frege Reader* (1997), pp. 151–171; also 'Logic' pp. 230–231, in the same volume; 'Introduction to Logic', in *Posthumous Writings* (1979), pp. 191–192; Alexius Meinong, 'The Theory of Objects', in *Realism and the Background of Phenomenology* (1981), pp. 76–117; Bertrand Russell, 'On Denoting' (1905), pp. 479–493; F. Berto and M. Jago (2019), *Impossible Worlds*, Ch. 11 pp. 239–299.

For inferentialism: Wilfred Sellars, 'Empiricism and the Philosophy of Mind' (1956), pp. 253–329; Robert Brandom, *Articulating Reasons: An Introduction to Inferentialism* (2000).

● SECONDARY READING

For recent treatments of **assertion**, the central reference is Timothy Williamson, 'Knowing and Asserting' (1996), pp. 489–523. The connection with epistemology can be explored further in his *Knowledge and Its Limits* (2000). An influential paper on these lines is Keith DeRose 'Assertion, Knowledge and Context' (2002), pp. 167–203. An excellent collection of recent work on the subject, including essays by the editors, is Jessica Brown and Herman Cappelen (eds), *Assertion: New Philosophical Essays* (2011). More recent is *The Oxford Handbook of Assertion*, edited by Sanford Goldberg (2000).

For fictional objects, the idea that fictional objects are possible objects is explored by David Lewis in 'Truth in Fiction', reprinted in *Philosophical Papers Volume I* (1983), pp. 261–275. Neo-Meinongianism is defended by Terence Parsons, 'A Meinongian Analysis of Fictional Objects' (1975), pp. 73–86, and *Nonexistent Objects* (1980). For the creationist view, see Saul Kripke [1973], *Reference and Existence* (2013). See also Anthony Everett and Thomas Hofweber (eds), *Empty Names, Fiction and the Puzzles of Non-existence* (2000), and Graham Priest, *Towards Non-Being: The Logic and Metaphysics of Intentionality* (2005).

For pro-context sensitivity, see Charles Travis, *Occasion-Sensitivity: Selected Essays* (2008), Avner Baz, *When Words Are Called For – In Defense of Ordinary Language Philosophy* (2012) and Anne Bezuidenhout, 'The Coherence of Contextualism' (2006), pp. 1–10. For anti-context sensitivity, see Herman Cappelen and Ernie Lepore, *Insensitive Semantics: A Defense of Semantic Minimalism and Speech-Act*

Pluralism (2005) and, more recently, Emma Borg, *Pursuing Meaning* (2012). For relativism as opposed to contextualism, see John MacFarlane, *Assessment Sensitivity: Relative Truth and Its Applications* (2014).

For inferentialism, see Jaroslav Peregrin, *Inferentialism: Why Rules Matter* (2013).

● NOTE

- 1 Further sources of purported context-sensitivity: (1) The case where a term occurs intransitively but is implicitly transitive, as in 'You're ready' being shorthand for 'You're ready to/for ϕ ', with ϕ being a term for some action or event. Suppose we are impatient to go to a party and you say 'Berta is not ready'; typically one can answer the question, 'For what?', but the sentence alone doesn't tell us what Berta isn't ready for. (2) Putnam argued that natural kind terms – like 'water' – contain a tacit indexical element, since what they refer to depends on the possible worlds in which they are used. As we saw when discussing two-dimensional semantics, it's plausible that utterances of 'water' at certain non-actual possible worlds express the same concept, or meaning, but the referent differs. In those contexts of utterance, the substance called 'water' would be, for example, XYZ. (3) Ordinary proper names can be argued to involve indexicality or to be indexicals themselves. To what person does one's utterance of the expression 'John Smith' refer? Answer: of the many John Smiths in existence, the utterance refers to the one at the origin of the causal chain of which one's use is the terminus. Thus we can represent the name as 'John Smith_x', where the parameter x selects the relevant causal chain.

14

• modern directions II

In this chapter our topics are slurs, conceptual engineering, and experimental philosophy (X-Phi).

● SLURS AND PEJORATIVES

Assume that by a ‘concept’ we mean the meaning of a word. **Slurs**, as we’ll use the term here, are pejorative words standardly used to disparage or harm people for being of a certain race, sex or gender, sexual orientation or other social group. ‘Wop’, ‘bitch’ or ‘fairy’ are examples (but not always; said of a person, ‘mousy’, ‘creep’ etc. are not like this but still pejorative). Slurs are plausibly a variety of **Thick Concepts** – ones that combine an evaluation with some non-evaluative description. *Genocide*, *littering* and *generous* seem to fit the mould of being thick concepts. Candidates for thin concepts include *knee*, *nitrogen* and *pencil*.

We must acknowledge immediately that mentioning the words rather than using them does not completely cancel their offensiveness – this is itself a telling fact about them – but despite some discomfort we will proceed in the spirit of doctors examining some nasty specimens of a disease. I’ve chosen ‘wop’ for my chief example because, though still familiar, it is somewhat out of date.

Our question is what distinctive sort of meaning – if it is meaning – do slurs have? What exactly is the point or significance of a slur?

(A) **Slurriness as augmenting truth-conditions.** A first thought is that the slurry-meaning is simply an ingredient of the *sense* or *content* of such words that is superadded to the sense or content of some non-pejorative, merely classificatory term. So ‘wop’ includes the sense of ‘Italian’ (or ‘of Italian descent’) but also includes further content, such as ‘is contemptible just for being Italian’.

The proposal runs into trouble because the term's content determines the term's extension. One may well take umbrage at the claim that the slur has the same extension as the non-slurring counterpart. Certainly, the thesis that the following is *true* is very controversial:

(1) All and only Italians are wops.

Some would say that most Italians are not wops. Others would say that *no one* is a wop, and still others would say that *necessarily* no one is a wop.

None of these seem entirely satisfactory. Furthermore, if slurs do their nasty work at the level of sense or content, then that nasty work should be capable of being *negated*. 'Non-wop' should undo the damage of 'wop'. If some racist says:

(2) A wop is coming to the party,

then it should be possible reject this speech-act including its nastiness by asserting:

(3) No, no wops are coming to the party.

But one asserting this, at least to many people, is still acquiescing in the use of a racist term, and thereby performs a racist speech-act. Likewise:

(4) If a wop is coming to the party, then I'll eat my hat.

This is evidently a racist thing to say, *even if the antecedent of this conditional is understood as not true*. The offensiveness of the term 'leaps out' of attempted quarantine by means of logical devices such as 'not' and 'if-then'. (In an alternative terminology, the contribution of the term 'scopes out' of negations or the antecedents of conditionals, and of similar devices.)

The slurriness also leaps out of attempted confinement in indirect speech reports – i.e. sentences of the form 'x said that S'. (5) seems at least slightly less objectionable than (6):

(5) Jolinda said: 'A wop is coming to the party'.

(6) Jolinda said that a wop is coming to the party.

It's the distinction between mention and use, between direct and indirect quotation. Many think that (6), unlike (5), inherits the racist overtones of (2), that a sayer of (6) is in cahoots with Jolinda as regards her attitudes to Italians.

The same thing appears to go on with propositional attitude reports, such as:

(7) Jolinda thinks that a wop is coming to the party.

– and other embedding constructions, such as questions, modals, conditionals and indirect speech and attitude reports.

(B) Slurriness as implicated or presupposed. These considerations also seem to rule out the obvious *pragmatic* explanations of slurs. If the derogatory content of a slur were a matter of Gricean conversational implicature, it would have to be *cancellable* (as discussed in Chapter 9, ‘Some applications of the concept of implicature’). One can cancel the implicature ‘I’m not going out with you tonight’ of ‘I’m washing my hair’ to a dinner invitation, by following it up with ‘But we can go later’. Likewise if implicature were all that was going on with slurs, one could non-offensively utter:

(8) A wop is coming to the party. But I dig them!

I assume (8) is as racist as (1) is, or at least almost as bad. You can see the attraction of explaining slurs in terms of conversational implicature: since the derogatory content is not said but implicated, then the criticism for using a slur would be that what one said may be true, but it is pragmatically improper. But the derogatory content cannot be cancelled, which it must, if slurs are to be subsumed under the category of conversational implicature.

The same problem afflicts the thought that racist attitudes are *presupposed* by the use of slurs, again in the technical sense of Chapter 9 (‘Presupposition: Strawson’s and Donnellan’s objections to Russell’s theory of descriptions’). The idea behind this strategy is again inviting. Conversations take place with certain presuppositions in place: from day following night, to one’s bank not having gone under while one slept, these items for presupposition need not normally be spoken of. Thus, demeaning but shared attitudes towards certain groups might be presupposed. However, if something is presupposed in a conversation – as in the case where we assume the truth of ‘We have beer in the refrigerator’ in discussing who is going to fetch it – it should be in our power to cancel that presupposition, as in ‘You know what? This conversation is pointless; there is no beer in the fridge!’ Yet it seems that the derogatory content of slurs, once they are present in a speech-act, *cannot* be erased. As if a skunk has sprayed, slurs cannot be taken back.¹

(C) Expressive meaning which is sharply distinct from their ‘semantic’ (truth-conditional) meaning. We have so far considered approaches to slurs which either add the slurring propositional content to the classifying element, or shunt off the slurring content to pragmatics. We turn now to some alternatives.

We begin with a brief outline of a treatment largely due to David Kaplan. As we will see, the solution is perhaps partial at best, but it does put front and centre an important point that has just emerged: the speech-act of uttering a slur seems typically to be an act like the act of giving the finger or other gesture, that once done cannot be erased. The trick is to dovetail that idea with the plain fact that slurs are part of language.

Go back to Frege. In addition to sense and reference, Frege spoke of the ‘tone’ or ‘colouring’ of words. This is the way he thought that ‘cur’ differs from ‘dog’: same reference, and *same sense*, but different in what he termed the subjective, psychological images evoked by their use. Likewise, the conjunctions ‘p and q’ and ‘p but q’

have the same sense but their tone differs: the latter, but not the former, prompts the listener to prepare for a contrast between the first conjunct and the second about to be voiced. So one might think that tone is a candidate for classifying slurriness, because tone, once in play due to the presence of the word itself, appears immune to attempts at quarantine – ‘If you bring that cur, then my cat will not be seen’ – seems to cast aspersions on the dog no less than ‘You’re bringing that cur!’

However, it’s doubtful whether the notion of tone is quite stable. It is a matter of convention that ‘She smokes and doesn’t drink’ differs from ‘She smokes but doesn’t drink’ in that the latter acknowledges that the expectations aroused by the first conjunct are thwarted by the second. But if it is a matter of *convention*, then that is just what it is to be a part of meaning. It is part of the objective use of language, not of the mere subjective world as Frege seemed to have envisaged.

Kaplan’s proposal recognises a category of conventional meaning that essentially ‘objectivises’ Fregean tone. In addition to their descriptive function – which covers reference and truth-conditions – there is for many words what Kaplan calls their *expressive* function or meaning. The expressive meaning of a word is a state or attitude *of the speaker* that is not descriptively conveyed – does not have any effect on truth-conditions – but is conventionally tied to the use of the term. For example, ‘That blasted Mr Smith is coming’ is descriptively equivalent to ‘Mr Smith is coming’, but they differ expressively.

Expressive meanings do *not* embed within truth-functional operators: ‘That blasted Mr Smith is coming’ and ‘It is not the case that that blasted Mr Smith is coming’ are contradictories but have the *same* expressive meaning. Expressive meaning leaps out of attempted confinement within the scopes of truth-functions and other types of embedding clauses (if you don’t share the speaker’s attitude towards Mr Smith, you’d say something like ‘Hey! Don’t call him that!’). This is a significant advantage over the appeals to implicature or presupposition just discussed.

Among the reasons that the theory is independently plausible is that it explains why expressions such as ‘Ouch!’, ‘Damn!’ or ‘Hello!’ cannot embed into sentences yet are nonetheless meaningful: they are descriptively meaningless – so they provide no items on which compositional rules of semantics can operate – but they are expressively meaningful. Yet they are conventional all the same, as is shown by different greetings and exclamations in different languages.

The descriptive function of ‘wop’ is the same as that of (the noun) ‘Italian’, on this account. The two are descriptively the same. But the expressive function differs. The expressive function of the word is to express hatred, disrespect or contempt for Italians. The illocutionary acts afforded by such terms are *acts* of insult or abuse, like the V-sign or the finger. This is why their use cannot literally be undone, any more than one’s actions yesterday can be erased from the past (which is not to say that their effects cannot be mitigated – by apologising, for example). That is why the word is not welcome: in normal circumstances, you cannot use it without expressing contempt for Italians. If you say ‘A wop is coming to the party’, you’re guilty of racism; if instead you say ‘No wops are coming to the party’, you’re still guilty of racism. It is a racist word.

More generally, the expressive function of epithets – ‘blasted Mr Smith’, ‘honourable Ms Jones’ – is to express attitudes of the speaker, real or feigned. Not that *expression*

of an attitude is the same thing as saying that one has it. For if the truth-condition of 'Mr Smith is coming' is the same as 'That blasted Mr Smith is coming', then neither can be equivalent to a statement that is explicitly about the speaker.

(D) Two-folded speech. Slurs are a little like platypuses, in that what might seem at first to be local, if disturbing peculiarities, prove on closer examination to be not only rich and revealing sources of data, but also as potentially threatening to certain widespread theoretical commitments. Accordingly, there are now on the market a multitude of views of slurs, many with the theoretical adjustments they appear to require of empirical semantics. In this penultimate sub-section, we consider very briefly two views which can be seen as responses to certain shortcomings of the broadly 'expressivist' view we've just been through. The two views – due to Elizabeth Camp and to Kent Bach – in effect take issue with the Kaplanesque claim that each slur has a univocal extension, even questioning whether, in typical cases, the truth or falsity of the corresponding proposition is well-taken.

Some have claimed that slurriness *can* be confined in a way that the approach we have just been considering rules out. Consider:

- (9) If his attitude towards Alessandro's coming to the party is that a wop is coming to the party, then sure as heck I'm not coming.

On a natural interpretation of this, it is held, the speaker cannot be charged with being responsible for the slur, but the slur is used and not merely mentioned. At least it is not mentioned in the straightforward sense of appearing inside quotation marks. This is a direct counterexample to the expressivist view of the preceding section, according to which the very function of slurs wherever they appear (except when merely mentioned) is precisely to express a certain demeaning attitude on the part of the speaker.

Partly in order to cope with such cases, Camp (2013, 2018) advances a 'dual act' theory of slurs. To use a slur is to perform a compound speech-act comprising an act of neutral description together with a commitment to the appropriateness of a certain *perspective*, one that derogates the object or objects of the slur. The expression of the neutral descriptivist element is, for example, what is common between 'wop' and 'Italian'. Unlike the expressive element of Kaplan, however, the ways in which perspectives figure into the understanding of particular uses of slurs depends much more on details of the context of use, and in some contexts the perspective represented by the slur can indeed figure in truth-conditions of the 'at-issue' content – the derogating propositions which are, in other cases, in the background, can be foregrounded. The notion of perspectives, meanwhile, is rich and detailed, but one has a rough and ready grasp of it: perspectives are intuitive, sometimes unwanted modes of interpretation that can be adopted by multiple agents, often as the result of social-historical construction, which guide and colour an agent's beliefs, attitudes and feelings towards the objects they present in an open-ended way.

As with the expressive account of the preceding section, the perspective is always in a sense present in the use of a slur, is *typically* not subject to being contained or 'quarantined' inside the scopes of indirect quotation and propositional attitude verbs

and the like, and as already intimated typically does not figure in calculating the truth-conditions of the 'at-issue' statement. But typical cases are by no means all cases.

First, as was evident if not emphasised in consideration of (2) and (3), we sometimes hesitate to allow that the uses of slurs are true when they express a reprehensible attitude (calling one 'true', one might think, is to *endorse* the statement, perhaps even to praise the speaker). On Camp's view, the indecision here is both real and justified. For the utterance comprises a dual-act involving two claims, which may diverge in truth-value. It gets 'something right *and* something wrong', Camp says. Whether or not to foreground one or the other depends on the 'focus of the conversation, the speaker's communicative purposes, and our own evaluative purposes and commitments' (Camp 2018, p. 51). As we see from (9), in some cases it is possible to understand the use of a slur correctly and to endorse it without buying into the relevant perspective. In other cases it is not possible, especially in the case of 'weaponised' uses of slurs, as when a speaker says unprompted to an Italian, 'You wop!' In both cases, we may rightly feel conflicted if we are asked for a final judgement, true or false.

Second, the varying capacity of slurs to break out of confinement within indirect speech or propositional attitude verbs and the like, can be explained. The perspectival element of a given slur varies enormously if subtly or even unawares in different contexts. Because perspectives are open-ended and amorphous background modes of interpretation, they cannot be encapsulated as a specific proposition or set of propositions which a hearer might accept or reject. Further, 'there is a default preference for predication of group membership to be at-issue and endorsement of a perspective to serve as commentary' (Camp 2018, p. 56). But if it is made abundantly clear that the attitude is not one that the speaker or any participant in the conversation shares – as in (9) above – then the attempt at quarantine may be successful. In such cases confinement occurs: 'it will bind the perspective', as Camp puts it. 'Binding does not itself *automatically* cancel commitment to the perspective by the actual speaker; rather, it merely enables the speaker to quarantine themselves from it *if* they sufficiently signal their distance' (2018, p. 51; emphasis added).

Camp's notion of a 'perspective' being the key to a theoretical understanding of slurs is compelling, but you might wonder whether the notion is sufficiently stable to play such an essential theoretical role. Kent Bach (2018) espouses a rather more straightforward conception of slurs – some would say at the price of sacrificing the more compelling subtleties of a view such as Camp's – that the use of a slur simply expresses two propositions, both forthrightly descriptive, rather than one being descriptive and the other perspectival or expressive in whatever sense. One is the neutral proposition ('x is Italian'), the other is a 'side-comment' ('x is contemptible just for being Italian'; 'contemptible' is meant as a generic stand-in, replaceable in practice with a similar negative property).

He dubs it 'loaded descriptivism'. Similar to Camp, since there are two separate statements to be considered, it is a mistake to suppose that a sentence like (1) has a single truth-value. But rather than context embeddedness determining the degree to which one proposition and hence its truth-value is salient, any such emphasis is regarded as a pragmatic feature of speech in general, not a feature specifically of

slurs. And Bach points out that such dual propositions are not unknown. Propositions of the form ‘*x*, who is *F*, is *G*’ are used to assert the two propositions ‘*x* is *F*’ and ‘*x* is *G*’. Slurs are distinct in this regard only due to their packaging two separable items of content into a single vehicle – e.g. ‘wop’ (and hence their power might at least partly be explained).

(E) Hom’s combinatorial externalism. The views we have discussed so far have it that slurriness involves a tandem relation, the slur part riding atop or aside the neutral counterpart, where the former is conceived as an attitude, perspective or opinion that typically if not always implicates the speaker, and perhaps the hearers as well. We can sum up the common part of such views as positing a sharp distinction between the descriptive component and what we can loosely call the evaluative component of a slur (allowing that the role may, as most explicitly in Bach’s notion of a ‘side-comment’, turn out to be straightforwardly propositional in character and therefore admit of truth-evaluability). To characterise this component of the slur in this way is an attempt to do justice to its visceral nature, to the discomfort or outrage one feels as a witness to its use.

Chris Hom (2008) advances a view which is conceptually tighter than the others we’ve been through, with a more integrated conception of their meaning, and whose price is that we dispense with the thought that a sentence which applies a slur to someone can ever be true. Thus we commit to ‘No Italians are wops’, and deny the implication of Kaplan’s idea that ‘A wop is coming’ can be true despite being evaluatively (expressively) noxious. It rejects the supposed tandem nature of slurs. Instead, the meaning of a slur invokes a certain social institution of racism (or other ‘ism’ for groups other than races; I will take this as understood for the duration). The content of Hom’s example of ‘ch—k’ is ‘ought to be subject to higher college admissions standards, excluded from managerial positions, because of being slanty-eyed, devious, all because of being Chinese’ (p. 431). The ‘combinatorial’ in combinatorial externalism is therefore ‘the view that racial epithets express complex, socially constructed, negative properties determined in virtue of standing in the appropriate external, causal connection with racist institutions’ (Ibid). To ascribe a slur is to assert that the target ‘ought to be subject to *these* discriminatory practices because of having *these* negative properties, all because of being NPC’ (Ibid.; NPC is the property attached to the slur’s neutral counterpart). A slur involves a complex property but it nevertheless is a single self-same one.

Some slurs are so extraordinarily powerful that one cannot even mention let alone use them outside special contexts. This feature is plausibly explained by the slur’s inevitably invoking the facts of racial discrimination and history, which vary tremendously from slur to slur. Perhaps the other pictures can accommodate the reference to the history of racism, but Hom puts it front and centre, no apologies.

It is a case of content ‘externalism’, since the precise nature of the meaning of a slur is not determined by the psychology of the speaker, but by the nature of the social institutions referred to, just as according to Putnam the semantical value of one’s use of ‘elm’ or ‘beech’ depends not on one’s narrowly conceived mental states, but

upon the botanist's classification, of which a competent and responsible user may be quite ignorant (see Chapter 8, pp. 141–144). One can be unable to specify a single property in which beeches differ from elms – except being called 'beeches' rather than 'elms' – yet nevertheless speak truly in saying 'That's a beech' and falsely in saying of the same tree 'That's an elm'.

To sum up, Hom claims the following advantages for combinatorial externalism:

- Slurs always have empty extensions. This contrasts with tandem views according to which one proposition expressed by a slur – the neutral proposition – is sometimes true.
- Non-derogatory uses of a slur pose no problem. Some examples: 'Alessandro is Italian, but he's not a wop'; 'There are many Italians in San Francisco, but no wops'; 'It used to be that many people in New York thought of Italian immigrants as wops' (examples adapted from Hom 2008). Slurs always express their normal derogatory semantic content, but sometimes they do not derogate anyone (just as one can use the word 'idiot' without calling anyone an idiot). Derogation occurs only when the derogatory content is actually applied, is predicated of a subject.
- Slurriness differs widely from slur to slur. Slurs for some groups are out-and-out vile and extremely demeaning. On the other hand, words like 'honky' or 'haole' are not, even if the connotation is negative. It is an historical and socio-economic matter to articulate why that is.
- Slurriness is independent of the attitudes of particular speakers, just as the meaning of 'beech' is independent of a given speaker's mental states. In either case, meaning is determined by social facts.
- Slurriness changes with changes in the social institutions and practices, sometimes to the point where the slurriness of the relevant slur lapses, as in the case of 'queer'.
- The phenomenon of reappropriation is a striking example of the last point. A counter-institution may arise 'to support the altered, appropriated use'. They do not 'wipe away' the derisive force of the slur: reappropriation involves the historical, derogatory semantic content, perhaps ironically, and somehow deploys it for other purposes (Hom, in conversation). They 'defuse' the original slur, and 'put it to alternative uses that produce political and social effects in favor of the previously targeted group' (p. 438). 'Queer' is an example; 'bitch' is perhaps another, and unlike 'queer', 'bitch' retains its status as a slur, even if its use is sometimes joking or ironic.

Hom's theory seems on the right track with its externalism: to call someone by a slur is a different and potentially graver thing than merely expressing one's subjective attitudes towards that person. It is to take part in a despicable social institution. But one may think that the theory does not quite do justice to feelings of divisiveness, of discomfort brought about with slurring speech, which were the motivation for the 'tandem' strategies of Bach, Camp and Kaplan. Kaplan's view at least, with its emphasis on expressive content, does some justice to the power of the words themselves, of their very phonology and morphology.

● CONCEPTUAL ENGINEERING

The phrase ‘conceptual engineering’ is a phrase redolent of optimism, suggesting modern teams of researchers bent on improving and revising our concepts, cleaning up conceptual messes and constructing new ways of thinking. Surely this is all to the good. Our concepts presumably constrain as well as make possible our beliefs, our hypotheses and perhaps indeed our desires. In one of his many prescient moments, Nietzsche wrote in 1901:

Philosophers ... have trusted in concepts as completely as they have mistrusted the senses: they have not stopped to consider that concepts and words are our inheritance from ages in which thinking was very modest and unclear ... philosophers ... must no longer accept concepts as a gift, nor merely purify and polish them, but first make and create them, present them and make them convincing. Hitherto one has generally trusted one’s concepts as if they were a wonderful dowry from some sort of wonderland: but they are, after all, the inheritance from our most remote, most foolish as well as most intelligent ancestors ... What is needed above all is an absolute skepticism toward all inherited concepts.

(Nietzsche 1968 [1901]: 220–221, §409)

Conceptual engineering may be motivated in various ways but principally the motives are theoretical/factual, logical/cognitive and moral/political.

The exact phrase ‘conceptual engineering’ seems to have been used only since Richard Creath (1990) used it, and the industry really got going only in the 2000s under that name. But the phrase refers to something that arguably has been around since the dawn of philosophy. What Plato meant by ‘forms’, or what Aristotle meant by ‘the golden mean’, at least one might suppose, are concepts which those figures thought up and introduced, that is, engineered. They might not be happy to have them called ‘concepts’, thinking that what they were concerned with was universals, was reality – not mere representational denizens of the mental realm – but for our purposes we can grant their point, and maintain that in addition to those entities of the realm of reference, there are also concepts, senses or the meanings of words, which enable us to think and communicate about those entities.

Besides the examples from antiquity, one can easily cite examples from the history of philosophy. Bentham excoriated the idea of ‘natural rights’ as ‘nonsense on stilts’ and introduced the idea of rights as established by political agreement. Hume, responding to the seeming incoherence of the common-sense idea of free will – as actions standing outside the causal nexus – can be thought of as having replaced that concept with one of actions which are not coerced by others. Marx introduced the notion of value as determined by labour. We saw in Chapter 7 Kripke positing his notion of a rigid designator. We saw in Chapter 9 Austin introducing the notion of illocutionary act. Outside philosophy narrowly so-called, one might think of examples from science (Newton and his definitions of force and momentum), from mathematics (the introduction of imaginary numbers, limits or the derivative),

examples from the law, from political science and economics. In psychology, perhaps Freud was acting as a conceptual engineer when he reified the notion of the unconscious, or those of the id, ego and superego. One might also cite examples which one runs across more frequently in one's ordinary life, such as *fish* (no longer applies to whales and dolphins); *planet* (no longer applies to Pluto); the latter was transparent and deliberate, the outcome of debate by international bodies.

Sometimes, we have *bad* conceptual engineering (just as we have the occasional bad engineering of bridges): the concept of the ether (a medium, everywhere present, through which light and gravity were thought to propagate), or the concept of phlogiston (a substance thought to inhere in flammable bodies, consumed in burning); science discarded these concepts as not having application, the pertinent phenomena being otherwise explained. Witches – women thought to have obtained magical powers from the devil in exchange for their souls – were once thought to walk the earth. The badness of bad conceptual engineering can be moral as well as social or scientific.

So clearly the fashioning of concepts has always gone on. Why then is conceptual engineering suddenly a thing? Two reasons come to mind. First, the phrase itself – itself an example of conceptual engineering – has had the practical effect of focusing minds on a particular cluster of issues (and the phrase makes this branch of philosophy out as something that is contemporary, active and collectively *done*, rather than as a passive and inert solo non-activity of navel-gazing, as philosophy is often taken to be). Second, much of the attention has been with the normative, and specifically the ethical dimensions of concepts – most visibly of the concept *woman*, for which Sally Haslanger (2000, 2012) has proposed a modernised alternative. The focus on conceptual engineering brings with it increased scrutiny on what readers of the section on slurs will recognise as 'thick concepts' – concepts which weld together their logical or classifying dimension with their evaluative dimension.

The remainder of this section will consist of: (A) discussion of a certain precursor to conceptual engineering known as conceptual analysis; (B) discussion of a historical halfway house to conceptual engineering known as explication; (C) more detailed discussion of the shape and content of conceptual engineering; (D) a focus on the ethical and political aspects of conceptual engineering.

(A) Conceptual analysis: a criticism. At the beginning of this book, we noted that philosophical questions are often of the 'What is X?' variety, for some philosophically interesting X. The answers to such questions are often incorporated into philosophical theories – *accounts* or *views* – of X. This is more or less what Socrates recommends in the Euthyphro. But the advent of the 'linguistic turn' in the first half of twentieth century – and for convenience we are pretending the movements of history to be simpler than they really are – brought with it a shift from talking about the things themselves or about kinds of thing, to talking about the meanings of words, about the concepts expressed by them – from X to 'X', from being to meaning, one might say.

It is natural to assume that a large measure of the linguistically inclined philosopher's interest is in conceptual *analysis*. Taking its cue from the analogy with chemical

analysis, the conceptual analyst seeks to elucidate the contents or Fregean senses actually expressed by philosophically important words. Once we know more about the concept, the clarity of the propositions formulated by its means is enhanced, the theories we enunciate by its means gain in sharpness.

The most basic form of analysis is of an explicit definition (familiar from Chapter 3): if C is the target – where for simplicity what is in view is the conceptual equivalent of a property rather than the conceptual equivalent of a relation – we seek something of the form

x is C if and only if x is ...

– where the material in the ellipsis is more fundamental or more articulate than C itself (we may wish to append the term ‘Necessarily’ to the whole thing, but we will not emphasise this). The two sides flanking ‘if and only if’ must express the *same* conceptual content (in Frege’s terms, they must express the same sense). A trivial example might be ‘x is a mare if and only if x is female and x is a horse’. A less trivial and more philosophically interesting example – and one that involves a relation-concept rather than a property-concept, a two-place predicate rather a one-place predicate – is ‘x knows that p if and only if p is true, x believes that p, and x is justified in believing p’.

The basic method, as exemplified by Socrates in the *Euthyphro* as well as in other dialogues, is to search for counterexamples: an exhaustive but failed search is good news for the proposed analysis. As a case in point, the proposal as regards knowledge just considered was once assumed without comment by most philosophers, but Edmund Gettier in his famous paper of 1963 pointed out some seeming counterexamples. Thus in general: examples to which C applies but which don’t fit the proposed definition, or which fit the definition but to which C does not apply, refute the proposed analysis. (The examples do not have to be actual examples; it suffices to describe a conceivable case which satisfies the one but not the other – a case which would, if it were actual, be a case of one but not the other.)

Conceptual analysis can take other forms. In addition to explicit definitions, it can take the form of a ‘contextual definition’ as described in Chapter 3; or an ‘implicit definition’: we identify a certain set of ‘conceptual roles’, of propositions which involve the target concept which are especially important to it $P_1 \dots P_i$, and declare them to be jointly definitive of the concept, ones that one has to know in order to understand the concept (a case: the so-called ‘**Peano Axioms**’ might be presented collectively as a definition of the concept *natural number*).

There are at any rate some serious objections to regarding conceptual analysis as characterising philosophical activity. The main trouble comes in the form of the **Paradox of Analysis** (it was the subject of question 4, Chapter 2). Suppose C is correctly analysed as P. If one grasps the sense or content of C, then one cannot obtain any information from ‘x is C if and only if x is P’, because of the requirement that C and P must have same sense or the same content. The analysis would be the same as the triviality that x is C if and only if x is C. Analysis surely is something more.

(B) Explication. There was a development in the mid-twentieth century that was plausibly a halfway house, or perhaps more than that, between conceptual analysis and conceptual engineering. Its most famous exponents were Carnap and Quine. What they called '**explication**' was motivated partly by concerns over vagueness. The envisaged improvements were gains in logic, clarity and ontology (the range of entities we have to acknowledge as existing).²

Many concepts we commonly use in ordinary life are too vague for scientific use. Although it is possible to think that this is a sort of illusion, that for every concept and every situation, there is a fact of the matter as to whether or not it applies, the apparent vagueness being only a shortcoming about our knowledge of concepts, most would agree that the uncertainty is located within the concepts themselves.³ In some situations, it appears that there is no fact of the matter whether a given object is large or not-large. This is seldom a problem in ordinary life. If added precision is required, we can often shift to the corresponding relational concept – in this case, to 'x larger than y' for 'x is large'. But as will emerge below, this ploy will not always be efficacious, especially not when matters of philosophical interest are involved.

Carnap saw the problem, writing 'in the case of an ordinary language ... the words have no clearly defined meaning' (1952, p. 427). Using them as they are often leads to 'sterile debates' (ibid), of people talking past each other unawares. Carnap cites the ordinary concept of probability as ambiguous as well as vague (three more precise concepts are given by the subjective, frequency and classical definitions): 'The task of explication,' writes Carnap 'consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second' (Carnap 1962, p. 3). Quine, echoing Carnap, said '[o]rdinary language is only loosely factual' (Quine 2008: 285). Commenting on the conundrums of personal identity, Quine writes:

They are questions about the concept of person, or the word 'person', which, like most words, goes vague in contexts where it has not been needed. When need does arise in hitherto unneeded contexts, we adopt a convention, or receive a disguised one from the Supreme Court.

(Quine 1996, p. 39)

Quine stressed that an explication must have a precise point or purpose, in the form of some accepted postulate, or conjunction of the same, that the newly fashioned concept must make out as true. Quine's prized illustration is the concept of an *ordered pair*. We know them from high school mathematics, but what exactly are they? Quine notes that the ordered pair is 'a device for treating objects two at a time as if we were treating objects of some sort one at a time' (1960, p. 257) and cites a 'typical use' of the device, that of assimilating relations to classes (sets). Consider the relation of *fatherhood*, which involves the pair comprising Abraham, Isaac (in that order). While order is immaterial in the case of some relations, in this case – the fatherhood relation – it is vital: in the usual notation, the pair <Abraham, Isaac> is a member

of our envisaged class whereas the pair $\langle \text{Isaac}, \text{Abraham} \rangle$ is not. The chief thing to recognise is that whatever the ordered pairs are, the following postulate must be true:

$\langle x, y \rangle = \langle w, z \rangle$ if and only if $x = w$ and $y = z$.

That is: one ordered pair is identical to another just in case their first members are identical and so are their second members. This postulate can be validated in various ways; an identification of the ordered pair $\langle x, y \rangle$ with the unordered pair of sets $\{\{x, \emptyset\}, \{y\}\}$ as Norbert Wiener proposes, or as $\{\{x, \{x, y\}\}\}$, as Kazimierz Kuratowski proposes, does the trick.⁴ Thus the notion of ordered pair does not require anything more than elementary set theory itself, which Quine is assuming.

Quine contrasts the answer given by Charles Sanders Peirce: the two individuals comprising an ordered pair is represented by a 'mental Dyad' consisting of two images with mentalistic symbols attached to them, one meaning 'First', the other 'Second'. Dismissing this as unhelpful, Quine declares: 'We do better to face the fact that "ordered pair" is (pending added conventions) a defective noun, not at home in all the questions and answers in which we are accustomed to imbed [*sic.*] terms at their full-fledged best' (1960, p. 258). We are free to add the missing conventions:

This construction is paradigmatic of what we are most typically up to when in a philosophical spirit we offer an 'analysis' or 'explication' of some hitherto inadequately formulated 'idea' or expression. We do not claim synonymy. We do not claim to make clear and explicit what the users of the unclear expression had unconsciously in mind all along. We do not expose hidden meanings, as the words 'analysis' and 'explication' would suggest; we supply lacks. We fix on the particular functions of the unclear expression that make it worth troubling about, and then devise a substitute, clear and couched in terms to our liking, that fills those functions. Beyond those conditions of partial agreement, dictated by our interests and purposes, any traits of the explicans come under the head of 'don't cares'.

(1960, p. 257)

Despite the words here about 'explication', Quine is agreeing with the main drift of Carnap's plan of explication. For a 'don't care' case, one can cite examples such as whether or not $\{x, y\}$ is a member of $\langle x, y \rangle$: if you're following Kuratowski then it is a member of it; if you're following Wiener then it is not a member of it, but no one cares about the answer one gives to the question.

Is conceptual explication not conceptual engineering? Essentially it is, but the purposes may be quite different. Neither Quine nor Carnap care especially if explicated words fail generally to 'catch on' (the 'implementation problem' as we'll shortly see); their recommendations are directed only at a small coterie of scientifically and ontologically minded philosophers. They would positively welcome it if philosophical problems lapse – if they cannot be formulated or have easy answers – using the explicated word (the 'trivialisation problem'); for them, this is what philosophical progress looks like. And most importantly, their main interest is in presenting a

streamlined version of science. Their notion of ‘improvement’ is purely logical and ontological, *viz.*, to present the facts of the world in the clearest and most economical way possible. Unlike those of most conceptual engineers nowadays, their aims are neither social, practical nor ethical.

(C) The shape of conceptual engineering. We shift now to the contemporary scene. One way to describe conceptual engineering is to describe it as subdividing into three tasks:

1. Notice some defective concept or non-existent concept, analyse and verbally articulate the problem that needs to be addressed, in whatever level of detail the problem requires, and make the case for a new concept.
2. Design a new concept, where this includes assessment of the prospects for logical, cognitive and moral/social improvement as well as the danger of unwanted effects. The introduction of the concept may take any of several forms, as outlined in the last section: explicit definition (normally a conjunction of characteristics), contextual definition (as in Russell’s definition of definite descriptions), implicit definition (providing a list of central statements involving the concept which may be presumed true) and others.
3. Implementation. In many cases this will be as much a political activity as a philosophical one: it’s a matter of getting people to use the new concept and to dispense with the old.

Many have distinguished between creating new concepts and fixing old ones, or between what David Chalmers (2020) has called De Novo engineering versus *re-engineering*. The latter divides into ‘heteronymous’ conceptual re-engineering (the new word is different), and homonymous conceptual re-engineering (the old word is retained but redefined). Such cases as Austin’s aforementioned *illocutionary act* can be regarded as De Novo. But most cases – especially in the cases of the ethically thick concepts discussed below – involve conceptual re-engineering, cases where faulty existing concepts are repaired. Both types will often depend on substantial amounts of work being done at Stage 1. Stage 2 will involve, not just judgements as to whether the proposal is ‘right’, but judgements as to whether the proposal would be useful, whether it would address the relevant needs identified at Stage 1 – be the needs logical, cognitive, scientific, normative, ethical or practical. Stage 3 is of the first importance, but it is a practical and political problem about which we will say little.

Related to this is the relatively important matter of what precisely one means by ‘concepts’, how concepts are ‘individuated’. For example, if some language lacked a single term for female sheep – the equivalent of the English term ‘ewe’ – but did have terms for females and sheep, would speakers of that language lack the concept *ewe*?

You could see the point of going either way. The case of ‘concept’ is itself a prime example of the sort of thing that Carnap would have thought an item ripe for explication. Maybe a person who has the concepts *female* and *sheep* thereby has the concept *female sheep*, and thus has the concept *ewe*. But on the other hand, we would probably hesitate to grant that a person who has the concept *multiplication* thereby has the concept *cube root*, just because the former is sufficient to define – in the simple sense

of an explicit definition – the latter. Similarly, we'd hesitate to grant that everyone before Austin already had the concept *illocutionary act*, since it is definable in terms they already had. There is a slippery slope there from the simple to the complex.

We shall proceed here on the somewhat artificial but simplifying assumption that *within* a language, to cite a concept is to cite the word *itself* – plus its meaning, where meaning comprehends not only the word's sense but its tone, colouring or rhetoric, its expressive meaning or even its perspective, to invoke Camp's discussion of slurs. This has the consequence that there are few if any concepts of French that are also concepts of English, but we can always speak loosely of concepts shared by French and English even if technically what we have are two equivalent concepts. Thus our *ewe*-benighted person, who has the ingredients for a definition of the equivalent English term 'ewe', does gain a new concept when they learn to call ewes by a single word. For us the very forms of words – their morphology, the way they sound and look – will be intrinsic to concepts. Concepts are flesh-and-blood artefacts of culture as much as cognitive abstractions. This will prove important when we get down to cases, making it plainly matter whether the proposal being considered is homonymous or heteronomous, in the sense introduced three paragraphs ago.

Finally, we should underline the point that conceptual engineering embraces the looser styles of definition as well as the tighter. We included above contextual and implicit definitions as earlier understood. We may go further, allowing styles that are looser still. We may go for Wittgenstein-inspired Family Resemblance methods of introducing new terms; we might say 'by a game we mean board games, games of solitaire, ball games, and things like that' – caring little for the apparent circularity of such statements, nor for their apparent indeterminacy. For the purpose is to bring the listener, by hook and crook, to an understanding of what is meant that is sharp enough for the envisaged purposes. We may even go for definition-by-paradigm, definition-by-stereotype or definition-by-prototype; for example, we might define 'leaf' by giving the example of an oak leaf, and saying 'and things of this kind'. This may be perfectly all right since the aim of conceptual engineering is often practical rather than buttoned-up logical; our methods may be frankly pragmatic, relying our listener's catching on.

It may well be that any method of definition – strict or loose – presupposes that the recipient, in some sense, already has the capacity to categorise things in the envisaged way. But the recipient has perhaps never thought of grouping them in that way. The value of a definition is that of enabling the recipient to actualise the capacity, getting them explicitly to notice the category and associate with it a label.

(D) Ethical, political and practical content. Consider concepts such as *woke*, *politically correct* or *real Americans*. In today's context, these forms of words have complex, politically charged and contestable meanings. The meanings have also proved changeable. 'Woke' was a sometimes-ironic feature of American black speech in the early 1960s (and perhaps before) that has since been appropriated for more general use by the political left, roughly expressing praise for how aware one is, and has lately come to be used pejoratively by the political right for certain features of the left. 'Politically correct' began as a gently humorous term in 1970s amongst the political left for what an unthinking outsider might think of as their Stalin-esque

tendencies; it has since become a term used primarily by the right pejoratively for virtually every position of the left.

Concepts are contestable for political reasons. Consider now the concepts *womanly*, *womanish*, *women's work* and *women*. At the minimum, it seems to be very much a question of context – the speaker including their intentions, the audience including their tendencies, the objects or persons of reference, and more – whether a characterisation of a person or activity, for example, as *womanly*, is a matter of praise or belittlement, and of the scope for misogyny in the associated speech-act. And, at least with respect to the first three, one can well get in the mood of regarding any use of these terms as inherently a political act.

Sally Haslanger (2000) has in effect pushed this thought with respect to the fourth case, the case of *woman*. There is in Haslanger's estimation a relatively inert classifying term *adult female human*, pertaining only to sex and age (though naturally one might wonder whether those can be shorn of political connotations), a near-biological classification. Haslanger's interest is in not sex so much as gender, and her task is not so much to describe as to prescribe:

... the task is not to explicate our ordinary concepts; nor is it to investigate the kind that we may or may not be tracking with our everyday conceptual apparatus; instead we begin by considering more fully the pragmatics of our talk employing the terms in question. What is the point of having these concepts? What cognitive or practical task do they *or should they* enable us to accomplish? Are they effective tools to accomplish our *legitimate* purposes; if not, what concepts would serve these purposes better?

(Haslanger 2000, p. 33)

The task is one of revisionary conceptual engineering, with a frankly political or moral purpose; 'neither ordinary usage nor empirical investigation is overriding', she says, 'for there is a stipulative element to the project: this is the phenomenon we need to be thinking about. Let the term in question refer to it.' The task is necessary not only because 'ordinary concepts are notoriously vague; individual conceptions and linguistic usage varies widely', but because 'inquiry often demonstrates that the ordinary concepts used initially to frame a project are not, as they stand, well-suited to the theoretical task at hand' (Haslanger 2000, p. 33). The task is to combat 'exploitation, marginalization, powerlessness, cultural imperialism, and systematic violence' against females; the task is transparently in aid of propaganda, but of laudatory sort. Thus she proposes a definition:

S is a woman iff S is systematically subordinated along some dimension *economic, political, legal, social*, etc., and S is 'marked' as a target for this treatment by observed or imagined bodily features presumed to be evidence of a female's biological role in reproduction.

Haslanger 2000, p. 39; this gives way to a more complicated definition at p. 42 but this simpler one will do for our purposes)

Haslanger is aware of the issue described three paragraphs from the end of the last section, of the power of the words themselves. Plainly, her proposal for 'woman' would be unlikely to have the desired effect if it were not a proposal for 'woman' but one for a new word, say 'gynexx'. The 'question of terminology is primarily a pragmatic and sometimes a political one', she writes; 'should we employ the terms of ordinary discourse to refer to our theoretical categories, or instead make up new terms? The issue of terminological appropriation is especially important, and especially sensitive, when the terms in question designate categories of social identity such as "race" and "gender"' (Haslanger 2000, pp. 34–35).

The radical nature of this instance of conceptual engineering is not just that it promises to empower one in the 'fight against injustice' (36); it stems from the very fact that it comes in the form of a *definition* (an explicit one). Thus, if the fight were successful – if all forms of sexual discrimination were put behind us – then, following the definition, females would lose their status as women. Haslanger celebrates this: 'it is part of the project of feminism to bring about a day when there are no more women', she pronounces (46); 'I'm happy to admit that there could be females who aren't women in the sense I've defined'. Any discomfort one might have at this consequence is a price worth paying, for:

... there is an invitation not only to revise one's understanding of these categories ... to revise one's relationship to their prescriptive force. By offering these analyses of our ordinary terms, I call upon us to reject what seemed to be positive social identities. I'm suggesting that we should work to undermine those forces that make being a man, a woman, or a member of a racialized group possible; we should refuse to be gendered man or woman, refuse to be raced.

(Haslanger 2000, p. 48)

If one feels discomfort at this – one may know, or be a female oneself, who is very much aware of sexual discrimination in all its forms, but who positively celebrates womanhood – then one can downgrade Haslanger's gender definition to a theoretical claim, a contingent categorical claim, one that is presumably true even if not without exception. One's struggle then would be to bring about a world in which the claim was false, in which women were not discriminated against.

If one does reply to Haslanger in this way, does this response expand into a criticism of politically ameliorative conceptual engineering in general? It threatens to, at any rate, if one intends homonymic conceptual re-engineering, and one wants to eradicate some undesirable property F from objects or individuals that fall under the target concept C, and one accordingly goes so far as to define C as possessing F.

The requirement accepted earlier, acknowledging the intrinsic power of words – that concepts be individuated partly according to the very words attached to them – comes into play in another way with respect to Haslanger's definition. Katherine Jenkins (2016) objects to Haslanger's definition of the concept *woman* on the grounds that it excludes at least some trans women (at the conceptual level, not just as a matter

of theory or fact). For, an appreciable number of trans women do not have – as they must according to Haslanger’s definition – ‘observed or imagined bodily features presumed to be evidence of a female’s biological role in reproduction’. They are excluded even if they are subordinated. Accordingly, Jenkins proposes a revision such that trans women are women. Matt Eklund in turn objects:

one may worry that insofar as the proposition regarding trans women that one seeks to defend is that trans women indeed are women, one does not defend that proposition by proposing a new concept of woman such that ‘trans women are women’ expresses a true proposition when ‘women’ expresses the new concept. To believe otherwise is tantamount to think that if we call a dog’s tail a leg, a dog will have five legs. To defend the original claim one must defend that trans women are women in the ordinary, pre-revision sense of ‘woman’.

(Eklund 2021, p. 20)

The complaint echoes a criticism of Strawson’s of Carnapian explication: Jenkins’ proposal merely ‘changes the subject’. But then one can feel that this misses the point: one is proposing a revision of the self-same gender-concept, precisely not to add a new concept. One goes along with the International Astronomical Union’s decision to revise the concept *planet*, and thus to change the status of Pluto from planet to dwarf planet, but the new practice is continuous with the old. One can do the same with the concept *woman*, except that the purpose is not astronomical but political. Jenkins’ proposal is motivated by feminist aims, not metaphysical aims.

The foregoing reasoning assumes that concepts are individuated partly by the words used to express them. Concepts have histories, exist in space and time, and their application conditions can change. Consider the concept *marriage*. Even if the notion of civil partnership were defined and accorded legal rights in exactly the same way as marriage, except the couple may be two females or two males, it will not be sufficient for same-sex couples who want to be married: the legal right to use the *word* is vital. Within limits, the concept can retain its identity through changes, just as the Ship of Theseus can withstand replacement of its parts while remaining the self-same ship.

● X-PHI (EXPERIMENTAL PHILOSOPHY)

(A) **Intuitions, conceptual analysis and reflective equilibrium.** Philosophy, including the philosophy of language, is full of arguments that rest upon ‘intuitions’. The word has multiple meanings, but here we mean a belief concerning some philosophically significant matter, for which one cannot provide further support, yet which seems right, even importantly right. Its denial is reckoned ‘counterintuitive’. A mainstream example will clarify this further. One has a ticket for a simple lottery; one winner out of 100,000 tickets (DeRose 1996). The draw has taken place, but the winner is yet to be announced. Does one *know* that one’s ticket will not be the winner? In spite of the very high probability that one’s ticket lost – higher than the probability

that your phone is your pocket, which nevertheless you do know – it seems that one does not know this. Perhaps one can't say why, but that is one's *intuition*. (After all, if one did know, then why would one buy a ticket in the first place?) One concludes that the high probability of being true is not itself sufficient for knowledge.

The lottery case fits the standard pattern of the 'Method of Thought Experiments', closely connected to the methodology of conceptual analysis discussed earlier, and what has been called in recent literature the 'Method of Cases'. This method fits with a larger method of philosophy, articulated by David Lewis, as one which seeks 'Reflective Equilibrium'. One has one's intuitions on the one hand, and one's theoretical commitments on the other. The two may be incompatible, where the theory is so entrenched or compelling that one recognises that one ought to doubt or withdraw the intuition. The goal in making the ensuing adjustments – of the entirety of statements one accepts, theoretical and intuitive – is a whole which, upon reflection, is the most satisfactory. Lewis writes:

Our 'intuitions' are simply opinions: our philosophical theories are the same. Some are commonsensical, some are sophisticated; some are particular; some general; some are more firmly held, some less. But they are all opinions, and a reasonable goal for a philosopher is to bring them into equilibrium.

(Lewis 1983: x–xi)

As we pointed out before, we can think of certain sorts of philosophical questions as about reality, or alternatively as about concepts. With a certain exception that will emerge, we will continue to think of such questions as ones concerning concepts. And we will continue to think of concepts as the meanings of words (but we can relax the last section's morphological requirement, the requirement that the word itself is necessary for the individuation of the concept).

(B) The Method of Surveys. Those behind the movement known as Experimental Philosophy, or 'X-Phi', pose certain problems for the above method, even where theory and intuition point the same way. Does the promulgator have the correct intuitions? How do they know that others have the same intuitions? The intuition may feel peculiarly compelling, but that is not going to persuade someone who has an opposed intuition. No scientist is likely to be impressed. 'So, it is your intuition that the world is flat? Hmm ...' And is the method of reflective equilibrium not unduly conservative? Is it not inherently unreceptive to the views and perspectives of others?

The proponent of X-Phi recommends instead that actual surveys be conducted. We settle on a certain way of presenting the thought-experiment or other question, and seek answers for some representative sample of the population. In this way, one's intuitions can be corroborated or defeated by the weight of public opinion.

Of course, the way to find out whether the earth is flat, for example, or whether dark matter exists, is not to carry out a survey (Kornblith 2002, 2013). Should one

think the same for questions concerning morality, or knowledge? (Here the matter is contentious; for example, it is not obvious that whether one should pull the lever in a Trolley Problem is a question about our concepts, rather than, simply, whether or not pulling the lever would be the right thing to do.) At any rate, *our* interest is not in facts of external reality but in facts about concepts, of the meanings of words. For this subject-matter – the contents of human thinking – what better than to ask human beings about their thoughts?

(C) Questions in the philosophy of language. Thus the case we shall centre on is itself a case explicitly about language. In Chapters 7 and 8 we discussed the ‘new’ or direct theory of reference. Such figures as Kripke and Putnam argued that contrary to the descriptivist views of Frege and Russell, for many words – proper names but also general terms for natural kinds such as ‘water’, ‘gold’ and ‘beech’ – reference is fixed not via an individuating description but via the practice of the surrounding linguistic community, or via certain causal chains of usage, either of which the user may have no conceptual grasp. We’ll focus on terms for natural kinds. The view in particular is that when one intends, for example, to call some liquid ‘water’, one invokes the scientific essence of the kind water, which is to be H_2O – even if one does not know what it is. Putnam in particular argued the case by using the results of a thought experiment: a person in 1700, faced with a sample of some non- H_2O substance XYZ, magically imported from another planet that mimics exactly the properties of H_2O , who called the alien substance ‘water’, would, strictly speaking, be making a mistake.

But are we merely to obey Putnam’s intuitions? (At least Putnam’s of 1975; he later backtracked from that position.) X-phi can help: any uncertainty over the Putnamian intuition can be resolved by putting the matter up for public vote. If only a distinct minority side with Putnam, that would be bad news for the new theory of reference. If a significant majority sides with Putnam, however, then that would be strong evidence that reference does work as Putnam says it does. For again, reference is a matter of the ways in which language is typically used; and the best way to find out how it is used is to ask the users.

Actual polling of direct relevance to the new theory of reference has emerged but, unfortunately, the results have been inconclusive (Jylkkä et al. 2009). But perhaps there is enough there to doubt the new theory; one thus wonders whether ordinary language really has terms for natural kinds at all. Or at least, one wonders whether anything like the model of reference proposed by Putnam is of the widespread significance that one might have thought.

(D) Problems with the X-Phi method. We shall mention very briefly some general issues thought to arise with X-Phi.

- Responses to questions on X-phi surveys appear to vary across demographic factors such as gender, ethnicity and social and economic class (e.g. Weinberg et al. 2001; Buckwalter and Stich 2013).

- Likewise for broader cultural factors. For example, differences have been reported between East Asian and Western students on epistemology and reference (Machery et al. 2004). Individual differences in verdicts about cases also appear to vary with personality type. For example, faced with questions pertaining to free will and responsibility, extroverts tend towards compatibilism (Feltz and Cokely 2009).
- However, evidence has emerged casting doubt on this 'Whorfian Hypothesis', that there are such significant differences (e.g. Nagel et al. 2013; Kim and Yuan 2015); subjects from different cultures have been found broadly to agree on Gettier-style thought-experiments (consistently with justified true belief not being sufficient for knowledge; Machery et al. 2015).
- Judgments about cases appear to be affected by matters of context that are irrelevant to the philosophical issue at hand. These include the manner of the presentation, the wording of the presentation and the order in which multiple cases are presented (as regards order, however, facing one case can provide evidence for a verdict on the next case, so the judgements may be affected quite appropriately; Horne and Livengood 2016).
- Some argue that the type of expertise enjoyed by professionally trained philosophers cannot be discounted; indeed, that judgements of ordinary people may be more or less irrelevant (Williamson 2007).

If one is impressed with the later Wittgenstein, one might object still more broadly to the whole idea of X-Phi, on the grounds that in making the same move as the traditional exponents of conceptual analysis – lending credence to judgements about described cases whether of the philosopher or of the subjects of experiment – it relies on a procedure that is plainly artificial. Stanley Cavell – and this 'radical contextualism' of his has since been taken up recently by Charles Travis and Avner Baz – points out that in actual conversation (the point is made in the first instance for moral conversation), the personalities involved must be considered, and also allowance must be made for the ins and outs of their 'moral relationship' (Cavell 1979, p. 284ff). This connects for example with the ordinary use of 'knows', but more generally with the method of cases. According to these figures, we too carelessly assume that underlying ordinary speech, concepts are sharp and unchanging, operating beneath explicit awareness. For Travis, Baz and the radical contextualist generally, it is sometimes quite unjustified to assume that a philosopher may expect an answer as to whether a concept applies to a case – in just those words in which the situation is described, independently of the fine details of the situation, and independently of the odd proclivities and tendencies of the participants, the specifics of their relationship and their interaction (Baz 2017, Travis 2008). Hence the need for literature, for not only Austin but Austen.

• CHAPTER SUMMARY

By *slurs* we understand thick concepts which disparage some social group – for example a race or gender. If we try to explain them in terms of their effects on

truth-conditions, we encounter the objection – if S is the slur and E a neutral classification of the target group – that the direct question whether ‘x is S iff x is E’ is true is an uncomfortable one. Similar problems arise with the idea that slurriness can be explained pragmatically, in terms of conversational implicature or presupposition. Both strategies require the possibility that a slur can be cancelled; it appears that they cannot be.

Kaplan developed (from Frege) the idea of a separate dimension of meaning from that which affects truth-conditions that he termed ‘expressive meaning’ (Frege: ‘tone’, or ‘colouring’). ‘Ouch’, for example, has no meaning relevant to truth-conditions but has a distinctive expressive meaning. ‘Tree’ contributes to truth-conditions but has no expressive meaning, or minimal expressive meaning. ‘Gentleman’ has both types, and slurs have both types as well. Plausibly, expressive meaning in speech-acts cannot normally be cancelled, and the strategy underwrites the thought that, again where S is a slur, ‘S is coming’ and ‘S is not coming’ are expressively equivalent, despite being each other’s negations.

However, some uses of slurs appear not to be themselves slurry. Cases where the slur is used – not merely mentioned – without slurring any person or group include cases of indirect speech, propositional attitude and hypothetical constructions. Camp and Bach propose tandem views of slurs. Camp’s is a subtle and flexible conception that combines an evaluatively neutral component with a *perspective*, a highly context-dependent feature which can in some contexts override their neutral component, thereby figuring in the determination of truth-conditions of the speech-act; but at other times it fades into the background, playing no role in the determination of truth-conditions. Bach’s view is straightforwardly that each use of a slur expresses two propositions, the neutral proposition and the evaluative proposition, a ‘side-comment’; the relative prominence of the two propositions from case to case is explained by more general pragmatic principles. Hom’s view is to bite the bullet, that all uses of slurs which predicate the slur of a person or group are in fact false; a slur’s content is explained as an objective matter of the history of prejudice against the relevant group, incurring the Kripke–Putnam externalist conception according to which one is responsible for the content of the term even if one is unaware of the relevant facts.

Conceptual engineering may be regarded as an alternative to conceptual analysis, which has been subject to objections, the most damaging of which is known as the ‘Paradox of Analysis’. Conceptual engineering is arguably as old as philosophy or nearly, but arguably came into its own only with the movement amongst philosophers of language deliberately to fasten onto concepts or the meanings of words, and self-consciously, for various explicitly acknowledged reasons, to change them. Notable incipient conceptual engineers were Carnap and Quine – Carnap in the 1930s, Quine most explicitly in 1960 – who called it ‘explication’, but their purposes were narrowly to replace vague ordinary concepts with precise ones, in the service of their narrow logical, scientific and ontological aims. More modern conceptual engineers broaden the idea, both in the logical kinds of engineering they pursue and in aiming for ethical and political change. Many of these theorists are explicitly concerned to

bring about social change by changing concepts, the conceptual lens through which we see the world.

The methodological question of whence the authority of the philosophers' 'intuition' is a sore one, and many do not find Lewis' answer satisfying of 'reflective equilibrium' – that intuitions are simply opinions, not sacrosanct, and philosophers must strive to bring their entire system including opinions and theoretical judgements into harmony. The replacement they recommend, for certain philosophically focused questions such as whether a subject in certain circumstances would have knowledge in a certain proposition, is to conduct surveys of ordinary people. If the philosopher does have an intuition with respect to the question, then the intuition may be corroborated or rejected in this way. This orientation – called 'Experimental Philosophy' or 'X-Phi' – is subject to some doubts. Perhaps philosophers do have some special expertise due to their training; perhaps the way a question is worded or the order in which questions are asked affects the results (matters of 'framing') and are of no relevance to the philosophical point at issue; perhaps answers vary depending on age, political or religious persuasion, gender, culture, personality type or geographical area; perhaps answers to such artificial philosophical questions yield a misleading picture of a subject's real-life propensities, manifested in their ordinary use of language in non-artificial situations.

• STUDY QUESTIONS

- 1 Suppose someone says: 'Words or concepts are politically inert; ameliorative conceptual engineering involves a confusion.' How should one respond?
- 2 Williamson thinks that just as physicists are experts with concepts of physics, so it is with philosophers and concepts such as knowledge. An example is the amount of training required to appreciate that knowledge is not justified true belief. More so when it comes to concepts of semantics – reference, sense, possibility etc. Does Williamson have a point?
- 3 A lecturer says, 'I see that there are no c... in the room.' There are no people in the room of Chinese extraction. One might say that the lecturer says something true, but that the conversational implicature is that c... *exist*, which is false. Is that satisfactory?
- 4 Consider the following inferences:
 Smith is coming. Therefore, that blasted Smith is coming.
 That blasted Smith is coming. Therefore, Smith is coming.
 Are these inferences correct? Think of validity as 'preservation of truth'; are they valid? What if you think of validity as 'adding no information'?
- 5 Dummett – with Brandom following – once suggested that slurring phenomena can be explained by inferentialism. For example, 'Boche' (an old racist term for Germans) might have its meaning explained by the following inference rules:
 'Boche' introduction: x is German; *infer* x is a Boche.
 'Boche' elimination: x is a Boche; *infer* x is barbarous and more prone to cruelty than other Europeans.

The case is supposed to be like ‘x is coloured’ being inferable from ‘x is green’. Is such a strategy plausible? Can it cope with the occurrence of ‘Boche’ within the scope of negation or in the antecedent of a conditional? Can such a story be told for other slurs?

● FURTHER READING (PRIMARY & SECONDARY)

For **slurs**, there is an edited volume: *Bad Words: Philosophical Perspectives on Slurs*, Sosa ed. (2018), which contains the relevant pieces by Camp and Bach (Hom’s original view was expressed in ‘The Semantics of Racial Epithets,’ 2008; the volume contains a later piece, co-authored with May, in which Hom explores some further issues). For constructive criticism of Hom from an inferentialist perspective see Diaz-Leon’s ‘Pejorative Terms and the Semantic Strategy’ (2019). Pope-Wyatt has written a comprehensive reader’s guide ‘Slurs, Pejoratives, and Hate Speech,’ for *Oxford Bibliographies On-Line*. Anderson and Lepore maintain that the badness of slurs is not to be explained in terms of ‘content’, even as broadly construed, but in terms of prohibitions: ‘Slurring Words’ (2013), pp. 25–48. Lepore’s more recent work with Stojnić extends the view, and promises to be a landmark: ‘Inescapable articulations: Vessels of lexical effects’ (2022).

For **Frege on tone or colouring** see the essay ‘The Thought’ [1918]; p. 155 of ‘On Sinn and Bedeutung’ [‘On Sense and Reference’ 1892], and p. 240 of ‘Logic’ [1897], all in *The Frege Reader*. For Kaplan see ‘The Meaning of Ouch and Oops’ (2004).

The amount of work in the past twenty years of so on **conceptual engineering** – both scrutinising the idea itself as well as promoting particular examples – has been enormous (see <https://www.youtube.com/@ConceptualEngineering>). Haslanger’s views on gender – ‘Gender, Race: (What) Are They? (What) Do We Want Them To Be?’ (2000) – have often been a point of reference, as here. For a focused and sustained view, see *Fixing Language: An Essay on Conceptual Engineering*, a monograph by Cappelen (2018). For a wide selection of views, see the volume *Conceptual Engineering and Conceptual Ethics* (2020); it contains the piece mentioned by Eklund, ‘Conceptual Engineering in Philosophy’; the view of Jenkins is in her ‘Amelioration and Inclusion: Gender Identity and the Concept of Woman’ (2016). Another useful collection is *The Routledge Handbook of Social and Political Philosophy of Language* (2021).

For the early precursor known as **explication**, see pp. 23–29 of Carnap’s *Logical Foundations of Probability* (1950, 1962), and Quine’s *Word and Object* (1960), §§33, 54. An early expression of doubt is in P. Strawson, ‘Carnap’s views on constructed systems versus natural languages in analytic philosophy’ (1963).

An even bigger industry – not only in Philosophy but in Psychology and Social Psychology – is **X-Phi** or Experimental Philosophy. See the entry ‘Experimental Philosophy’ in the *Stanford Encyclopedia of Philosophy* (<https://plato.stanford.edu/entries/experimental-philosophy>); monographs include Deutsch, *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method* (2015), and

Knobe and Nichols, *Experimental Philosophy* (2008). For ordinary judgements about Gettier-style cases see Machery et al., 'Gettier Across Cultures' (2015); Kim and Yuan, 'No Cross-Cultural Differences in the Gettier Car Case Intuition: a Replication Study of Weinberg et al. 2001' (2015); and Myers-Schulz and Schwitzgebel, 'Knowing that P Without Believing that P' (2013). For questions pertaining to reference etc., see Machery et. al., 'Semantics, Cross-Cultural Style' (2004); Deutsch, 'Experimental Philosophy and the Theory of Reference' (2009).

For defence of the philosophers' intuition against mass opinion, see Williamson, *The Philosophy of Philosophy* (2007). For general doubts about X-Phi, see Horne and Livengood, 'Ordering Effects, Updating Effects, and the Specter of Global Skepticism' (2017); Schwitzgebel and Cushman, 'Philosophers' biased judgments persist despite training, expertise and reflection' (2015); Feltz and Cokely, 'Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism.' (2009); Brown, 'Experimental-philosophy, contextualism and subject-sensitive invariantism' (2013); DeRose 'Contextualism, Contrastivism, and X-Phi Surveys' (2011). The radical contextualist charge against X-Phi, that holds equally against the appeal to philosophical intuitions, is, in effect, in Cavell, *The Claim of Reason* (1979), and in Diamond, 'Losing Your Concepts' (1988); more explicitly in Baz, *The Crisis of Method in Contemporary Analytic Philosophy* (2017), and Travis, *Occasion-Sensitivity: Selected Essays* (2008).

● NOTES

- 1 Attempts have been made to class slurs as 'conventional' implicatures as described Chapter 9, 'Implicature', fn 1. However, it is controversial whether conventional implicatures exist. Though it does have its adherents, it is not a popular strategy to invoke with respect to slurs.
- 2 The question of whether the two figures agree on this, and in what sense, has been the subject of detailed scholarly debate. For our purposes their rough agreement is more important.
- 3 One alternative is the epistemic account of vagueness, as a form of ignorance (so although we cannot normally know which, there is some exact number of grains of sand such that removing a grain of sand turns a heap into a non-heap), as in Williamson (1994).
- 4 'Nearly enough', writes Quine (1960: 257); Weiner's actual definition is more complex, but the philosophical point is the same.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

• glossary

acquaintance

See **knowledge by acquaintance**

anaphoric

Of pronouns and indexicals; an anaphoric use of a term ‘borrows’ its referent from some other referring device in the sentence (or beyond; it may borrow from other sentences in the context). For example, the indefinite pronoun ‘it’ in ‘Bob loved the symphony, but Billy hated it’; one could replace ‘it’ with another use of ‘the symphony’ without altering the meaning of the sentence (such have been called ‘lazy’ pronouns). By contrast, in ‘Whatever she wants, you’ll buy it for her’, the role of ‘it’ is essential to the meaning of the sentence (it is the role of a quantificational **variable**); and in ‘I am ready!’, the pronoun refers indexically, not anaphorically.

assertion, assertoric

Of sentences, the mood one uses to *say* something, make a statement, advance a proposition as true; it is primarily the indicative **mood**, as opposed to the nonassertoric moods such as the interrogative (questioning), optative (wishing) or imperative (commanding) mood. Of **speech-acts**, it is the act of making a statement, of verbally committing oneself to the truth of a proposition.

atomic proposition

The **proposition** expressed by an **atomic sentence**. Also referred to as ‘singular proposition’, and as ‘elementary proposition’.

atomic sentence

A sentence containing only one n -place **pure predicate** plus n singular terms.

attributive use

See **referential use**

biconditionality

A *conditional* is a statement of the form ‘If P then Q’; a *biconditional* is of the form ‘P if and only if Q’, which is the same as ‘If P then Q, and if Q then P’. It means that P and Q have the same truth-value: either both are true or both are false.

character

(Kaplan) Together with the **context of utterance**, the character is a rule which yields the **content** of a given expression, which in the case of a sentence is the

proposition expressed. The character of 'I am beautiful' yields different propositions depending on the identity of the speaker; analogously for other indexicals.

circumstance of evaluation

(Kaplan) Given the proposition expressed by a sentence at a given context, we can ask for an evaluation of it as true or false in the actual world or in non-actual but possible worlds. Similarly, we can ask for the evaluation of definite descriptions; at different circumstances of evaluation, different objects may be denoted by a given definite description. Once the **context of utterance** is determined, the reference of an indexical remains unchanged in different circumstances of evaluation.

cognitive content, cognitive meaning

As opposed to the expressive content or expressive meaning of a sentence, **singular term** or **predicate**. It is that element of content that is relevant to truth-conditions and reference.

cognitive value

Also called information value. One can learn something from 'The author of *Middlemarch* is George Eliot' that one cannot learn from 'George Eliot is George Eliot'; Frege called it a difference in cognitive value, which he explained as a difference in **sense** between 'The author of *Middlemarch*' and 'George Eliot'.

competence

As in linguistic competence. The ability, which may go unrealised, to say and understand linguistic items. A person may be unable to speak or even understand – due for example a brain injury – but the person's competence is unaffected. To be distinguished from language *performance*, events of actual speech or understanding.

compositionality

There is no upper limit to the number of grammatically correct sentences of a natural language such as English. Furthermore, we understand and produce sentences which are new to us. We are thus led to posit the principle of compositionality: that a finite vocabulary together with finitely many rules of grammar engender every possible grammatically correct sentence. ('An infinite capacity from finite means'.)

compound

See **indexicals**

concept

In psychology, by a 'concept' one typically means a type of cognitive capacity such as the ability to recognise *flowers* or *the colour red*. It is not normally thought to be tied to any particular language. Frege used the word somewhat eccentrically as the *referent* of a predicate, where predicates are interchangeable so long as their extensions are the same. So 'is a ten-foot-tall human' and 'was manufacturing transistors at the time of Julius Caesar' would designate the same concept, since their extensions are the empty set. Much closer to ordinary use, though not exactly the same as the psychological use, is to let concepts be the *senses* of predicates, as according to Alonzo Church. It's also common to speak of concepts as the 'content' of words.

conceptual engineering

It is significantly analogous to **explication**, but with a much greater emphasis on aspects of meaning of terms, besides those relevant to the truth-conditions of sentences: rhetoric, political and moral dimensions, tone, colouring and spin. It is a large and diverse movement in contemporary philosophy, which makes the topic difficult to pin down, but it is built into the words or concepts themselves that language, as we have it, can shape our thinking, and not always in cognitively and ethically healthy ways; therefore engineering is called for.

conditionality

In this book, the relation signified by ' \rightarrow ' between any two sentences P and Q, such that if P is true but Q false, then ' $P \rightarrow Q$ ' is false, and otherwise ' $P \rightarrow Q$ ' is true. This is one more precise replacement (**explication**) of the 'if-then' expression of English, whose interpretation is contentious. See **counterfactual conditional**.

conjunction

The conjunction of two sentences is true if both of those sentences are true; otherwise false. Typically formed in English by means of the word 'and'. There are also conjunctive noun-phrases.

connote

The descriptive content of a term; normally contrasted with **denotation**, which is the entities constituting the extension of the term. General terms – 'dog', 'sticky' – are generally considered to have connotation, but that all singular terms have it is controversial; many people following Kripke maintain Mill's doctrine that proper names only denote objects and do not have connotation.

content

Normally short for the cognitive content of an expression at a context. The term is often used in place of 'meaning' in the philosophy of mind and cognitive science.

context of utterance

The circumstance in which an expression is uttered or might be uttered; it includes time, place, identity of speaker, listeners and items demonstrated. See **circumstance of evaluation**.

context-relativity

Many words shift their reference depending on the circumstances of their use; this is obvious in the case of **indexicals** such as 'here', less obvious in the case of words such as 'tall'. Some philosophers hold that context-relativity is rampant, infecting even philosophically important words such as 'knowledge', 'permissible' and 'necessarily'.

context-sensitivity

see **context-relativity**

contextual definition

Unlike explicit definition, a contextual definition of a sub-sentential expression does not provide an equivalent expression, but provides a rule for converting whole sentences containing the defined expression to equivalent sentences that

do not contain the defined expression. The classic example is Russell's contextual definition of definite descriptions: The F is $G =_{df}$ there is an F , there is not more than one F , and all F are G . No expression equivalent to 'the F ' appears on the right-hand side of the definition.

contingent

A proposition that is true but not necessarily true is a contingent truth; a proposition that is false but might have been true is contingently false.

conventional implicature

A 'non-cancellable' variety of implicature – meaning that one is always responsible for it and cannot undo the implication– but still it is not a matter of formal logic. For example, 'The dog is very old but eats well' has the same truth-condition as 'The dog is very old and eats well' but is unlike it in that it conventionally implies some proposition as that the dog's being so old makes it unlikely to eat well. Contrasts with **conversational implicature**, which is cancellable.

conversational implicature

A statement that is not logically entailed by a given statement, but which is such that a reasonable and well-informed person (see **cooperative principle**) would take the statement to have been intended to be conveyed by the given statement. The implication is cancellable: the conversational implicature of one's saying 'I'm busy' in response to a dinner invitation might be a refusal, but one may undo the refusal by saying 'But bollocks, let's go!'

cooperative principle

A general constitutive norm governing conversation which comprises several sub-norms, including: maxim of quality (do not say what you believe to be false; and do not say anything for which you lack adequate evidence), maxim of manner (be brief; and avoid obscurity, ambiguity, vagueness), maxim of relation (be relevant) and maxim of quantity (make what you say neither more nor less informative than is required for the purposes at hand).

co-referential

Of two terms which refer to or stand for the same thing – for example, 'Norma Jean Baker' and 'Marilyn Monroe', or 'The first man to walk on the moon' and 'Neil Armstrong'. Co-referential terms may have different meanings or senses.

counterfactual conditionals

A conditional whose antecedent is understood as false. Normally expressed in the subjunctive mood – for example, 'If you *had scored* the goal, we *would have won*', said to someone who missed the goal. A popular analysis of 'If P had been, then Q would have been' is that in the possible world most similar to the actual world at which P is the case, Q is the case.

criterion

This word is used in various ways. The most famous user of the word was Wittgenstein, and his use of it has been the subject of considerable dispute. Roughly, to say the predicate B is a criterion of the predicate A is to say that facility with

A depends upon facility with B, where B is in some sense more basic than A, and where it is at least a *prima facie* rule that either (a) nothing can satisfy A without satisfying B (being B is a *necessary* condition of being A) or (b) everything satisfying B satisfies A (being B is a *sufficient* condition for being A).

declarative or indicative mood

The mood of a sentence normally used to express a commitment to something as true, as opposed to the imperative or interrogative moods. See **assertoric** and **force**.

de dicto

see *de re*

definite description

Using a noun phrase such as 'book on the table', it is possible to form an expression by attaching the word 'the' to the front, i.e. 'the book on the table', which can, in turn, be used in a sentence such as 'The book on the table is yours', which is true just when the predicate 'is yours' is true of a single book on the table. It is debatable whether the definite description 'the book on the table' is a singular term, along with proper names as Frege thought, or whether, with Russell, we should think of it as a quantifier phrase, like 'every book on the table'. See **denote**, **contextual definition**.

demonstration

A gesture such as pointing used to secure the referent of the use of a **demonstrative** (a type of **indexical**).

demonstratives

Indexical expressions such as 'this' or 'that bird', which require a **demonstration** to secure a referent. The demonstration may be implicit, if the intended referent is salient in the context.

denote

Normally a synonym for 'refers to'; sometimes used for 'mean'. Russell used it technically for the semantics of **definite descriptions**, for the relation between 'the F' and the F, when the unique F exists. This is *not* reference, in Russell's view, because reference is a pre-condition of the meaning of expressions; 'The F is G' is meaningful – indeed false – even when there is no such thing as the F.

de re

'Of the thing', in statements of modality (necessity, possibility etc.) and propositional attitude. It is contrasted with *de dicto* ('of what is said') in their behaviour with respect to **substitutivity**, the substitution of identicals. Consider the *de re* 'Of the Uffizi, John believes that it is in Rome' (more colloquially: 'You know the Uffizi? John thinks it's in Rome'). Together with the premise 'Uffizi = Florence's most famous art gallery', such a statement logically implies 'Of Florence's most famous art gallery, John believes that it is in Rome'. Crucially, the term outside the belief-operator or predicate – 'the Uffizi' or 'Florence's most famous art gallery' – need not reflect the subject of the ascription's own way of

thinking of its referent. The *de dicto* 'John believes that the Uffizi is in Rome' does not together with that premise imply 'John believes that Florence's most famous art gallery is in Rome'. 'John believes that the Uffizi is in Rome' can be true where 'John believes that Florence's most famous art gallery is in Rome' is false.

D-structure

'Deep structure'. In Chomsky's Universal Grammar (UG), roughly, a level of grammatical representation of a sentence in advance of any movement or grammatical transformations. See S-structure.

determines

In a Fregean framework, the relation between **sense** and **reference** of an expression. For example, the sense of the term 'The first man to walk on the moon' determines the referent of that phrase, Neil Armstrong. More generally, the components of a typical sentence refer to various entities; the corresponding senses of the parts of the sentence determine those entities as referents. The sense of the entire sentence, the proposition expressed by it, determines a truth-value.

directly referring expressions

An expression that refers to an entity without importing a Fregean sense, a definite description, a descriptive condition or a conceptual representation into the propositions expressed by sentences containing the expression. They **denote** objects without **connoting** anything.

disjunction

The *inclusive* disjunction of two sentences is false if both of those sentences are false; otherwise true. The *exclusive* disjunction of two sentences is true if exactly one of those sentences is true; otherwise false. In this book, we generally use the inclusive variety, as is typically formed in English by means of the word 'or'. The exclusive sense is often signalled with 'either ... or ...'

division of linguistic labour

We sometimes use words in a way that relies on the linguistic knowledge of others. For example, if one did not know anything that distinguishes beeches from elms (apart from their names), one could nevertheless speak falsely if one said, of a beech, 'It's an elm', and speak truly, if one said 'It's a beech'. Similarly, if one were to say 'It's gold!' of a sample of fool's gold, even if one lacks the knowledge that distinguishes the one from the other. According to Putnam, these capacities show that at least in some cases one's linguistic abilities rest upon the knowledge of others.

elementary proposition

See **atomic proposition**.

empty singular terms

A singular term that fails to denote or refer to an object, such as the proper name 'Pegasus' or 'Vulcan', or the definite description 'The man who walked on the planet Mercury'. Empty singular terms are not precluded from having cognitive meaning; Pegasus does not exist, but still we can talk meaningfully using the name.

epistemology/ical

Epistemology, or the *theory of knowledge*, comprises, among other things, theories of perception and of the justification of belief, and of knowledge itself. See **metaphysical**.

events, event semantics

Davidsonian semantics quantifies over events, rendering for example 'John ran' as 'There is an *e* such that *e* was a running and the agent of *e* was John'; this makes it possible to capture facts about adverbs, such as that 'John ran' follows from 'John ran quickly'. The question of exactly what events are, and of their 'ontological category', is very much a live one (beyond their distinguishing mark that they *happen*). If an event is intentional, then, arguably, it is an *action*; another large topic in metaphysics and semantics, the *theory of action*.

existential quantifier

'There exists a ___' is an existential quantifier. 'Some', 'something', 'someone', 'a' and others are not always used as existential quantifiers – they also have uses as **universal quantifiers**. It is complicated to explain in any detail the rules for when the expression is an existential as opposed to universal quantifier. But the distinction itself is straightforward. The existential – 'There exists an *F*' or 'Something is *F*' – is true if and only if at least one *F* exists. A universal quantification epitomised by 'Everything is *F*' is true if and only if everything is *F*. Other quantifiers include 'At least three things are *F*' and 'Most things are *F*'.

explication

Some philosophically (or scientifically) important words are unacceptably vague. Explication of such words is their replacement – where the replacements often sound and look the same as the old ones and which include the main part of the extension of the old ones – but where the new ones are sharper and more precise in theoretically relevant ways. Typically, where the term is a general term, the explicated terms will have more definite extensions, or will align with other theoretical needs in a more satisfactory way. An easy example is 'fish': the explicated term no longer applies to dolphins, but overall the term is rendered more useful for biology. Avoids the **paradox of analysis**; a precursor to **conceptual engineering**.

explicit or direct definition

The simplest kind of definition: an expression of a given semantical type is given a definition in terms of another expression of the same semantical type. For example: 'x is a mare =_{df} x is a female horse'. The classical line is that such definitions are cognitively eliminative; they are very useful in practice, but in principle, they do not add anything to cognition since, for example, one could always dispense with the word 'mare', using the expression 'female horse' instead. Contrasted with **contextual definition**. See **Paradox of Analysis**.

expressive meaning

A dimension of meaning, contrasted with **cognitive meaning** or descriptive meaning, which does not play a role in determining the sense, reference or truth-conditions of expressions. If 'bunny' and 'rabbit' are cognitively equivalent, then the difference between them is a difference in expressive meaning. What they express is

typically some attitude of the speaker. The concept makes sense of expletives and the like. For example, it's plausible to say that 'Damn!' does not have a truth-condition and does not refer to anything; but it serves to express a state of the speaker.

extension

Of predicates, the set of things that **satisfy** the predicate (the set of things which the predicate is *true of*); of a singular term, it is the term's referent; of sentences, it is the truth-value of the sentence or the set of worlds at which the sentence is true (all of this has also to be relative to contexts of utterance).

extensional

Of a language, it means that in any sentence, a singular term, predicate or contained sentence may be replaced with another of the same **extension**, and the result will have the same truth-value as the original. Individual positions within sentences may also be called extensional. Natural languages such as English are not normally thought of as extensional; **non-extensional** positions include those within sentences ascribing **propositional attitudes** and those containing modal operators, *viz.* 'possibly' and 'necessarily'.

external language

See **internal language**

external negation

'It is not the case that b is F' is an *external* negation, as opposed to 'b is not-F', which is an *internal* **negation**. On some accounts of singular terms, if b does not exist, then the external negation is true but not the internal negation.

family resemblance

According to Wittgenstein, many terms – including the term 'language' – pertain to things not in virtue of their essence, or necessary and sufficient conditions for the term's application, but through various overlapping resemblances. There is no property necessary to a thing's being a *game*, for instance. Wittgenstein adds that such terms express 'open concepts' – that is, the family of resemblances indicated by the term is constantly subject to revision.

fictional objects

'The Golden Mountain', 'Elizabeth Bennett', 'Batman'; these are all singular terms but they are different from 'Richard Nixon', 'The Eiffel Tower' and 'Antarctica'. Those of the first lot are not real things like the those of the second lot; many would say that those of the first lot *don't exist*. How to explain the first lot is the problem of fictional objects.

fix the reference

According to Kripke, a definite description or indexical may be used to introduce a referring term by specifying its reference – 'fixing it' – without thereby being understood as synonymous with it, as having the same sense or cognitive content as the term.

force

The type of speech act that one performs in uttering an expression (in Austin/Searle terminology, the *illocutionary act*). The most basic varieties are assertoric, interrogative and imperative. Closely but not invariably connected is the corresponding **mood** of the sentence used, which is a matter of grammar. For example, normally one asks a question by using the interrogative-mood sentence 'Are you shining my shoes?', but one can also do it by using the indicative-mood sentence – a sentence that normally is used for assertion but in this case is used with a special intonation: 'You're shining my shoes?'

forms of life

Wittgenstein's term for the various kinds of human activity in which the use of language is often embedded. Greeting, building, sport, sitting in the seminar room, haggling, ordering food, banter ... there are many forms of life, each requiring distinctive skills and know-how.

general term

Terms like 'dog', 'red', 'smokes' and 'kissed', which either (1) combine with the 'is' of predication and a term such as 'a' to form a **predicate** such as '___ is a dog'; (2) combine with the 'is' of predication to form a predicate such as '___ is red'; or (3) form a predicate without the benefit of additional words, such as '___ smokes' or '___ kissed ___'. See **singular term**, **definite description**.

grammar

A grammar in the modern sense for a language is a theory which includes the lexicon of the language (roughly its vocabulary), and a set of recursive rules which determine all and only the well-formed 'strings' of the language (its sentences etc.). A traditional grammar or 'school grammar' is implicitly the same thing, but not nearly as precise as a grammar promoted by Zellig Harris or Noam Chomsky.

historical chain

Kripke's term for a sequence of reference-preserving links of a term that connect one's present use to the origin of the chain, normally an act of dubbing. What preserves the link is the individual implicit intentions of each user, or perhaps a norm held by the linguistic community.

holism

Suppose we consider a set of objects and some property that each object in the set is a candidate for bearing. To say that the property is *atomistic* is to say that each case can be decided definitively without reference to the other cases; to say that the property is *holistic* is to say that none can be decided definitively unless all cases are decided.

hyper-intensional

A non-**extensional** context in a sentence, which does not guarantee substitution by **co-referentials** without changing the sentence's truth-value, unless they have the

same **cognitive meaning**. A less demanding type of non-extensional context than a hyper-intensional context is an **intensional context**.

illocutionary act

In the Austin/Searle scheme, an aspect of a speech-act that is typically intentional, and determines the illocutionary **force** of the utterance. For example, a particular kind of illocutionary act is that of asking a question. It contrasts with the **locutionary act** of expressing a certain content or proposition, and the **perlocutionary act** of getting the audience to answer.

illocutionary silencing

Sometimes one fails to perform the **perlocutionary act** that one intends. For example, you say 'Stop that!' to a child that nevertheless keeps on shooting the cat with a squirt gun. Many think that one can likewise find oneself unable to perform a **locutionary act** due to 'uptake failure'; these are cases of **locutionary silencing**. For example, a child might not be counted by the participants as genuinely making a contribution to an adult conversation, even though the child uttered a sentence which was clearly heard by the participants. Or to take a famous example, in a highly sexist environment, a woman might find that what she intends to be her refusal to have sex is taken as acquiescence (the relevant males say: "When she says 'no', she means 'yes'").

implicature

See **conversational implicature** and **conventional implicature**. Not to be confused with *implies*, or with *implication*, which is a matter not of pragmatics but of logic; to say that one statement implies another is to say that the one logically entails the other.

implicit relativity

Many expressions are used as if they denoted properties, but actually denote relations, with one relatum tacit. For example, 'The shirt fits' is really 'The shirt fits A', where A is some person.

indexicals

Expressions whose reference varies systematically with the context of utterance – shifts in time, place and possible world. The standard line is that 'here', 'then', 'now' and so on are *pure* indexicals (or have a standard use as pure indexicals). Indexicals that are not pure indexicals – called demonstratives – require objects to be demonstrated; examples include 'it', 'that', 'this', 'those', 'them' and 'there'. See **context relativity**.

indicative mood

see **declarative mood**

inferentialism

A theory of meaning which takes pragmatics – and in particular the speech-act of inferring – as fundamental and explains semantics in terms of it.

intension/intensional context

The *intension* of an expression determines the referent (or extension) of the expression at each possible world. 'Jane Austen wrote *Sense and Sensibility*' and

'Jane Austen wrote *Pride and Prejudice*' are both true in the actual world, have the same actual extension, but in some worlds have different truth-values; hence they have different intensions. An intensional position or context is one such that the substitution of a co-extensive expression may affect the extension of the containing expression. Only if the substituted expression has the same intension as the original does the substitution preserve truth-value across all worlds. Intensionality is not as fine-grained as meaning itself; for example, 'Any kangaroo is a kangaroo' and ' $2 + 2 = 4$ ' have the same truth-value in every possible world – they are always true – but don't mean the same. See **extensional**, **hyper-intensional**.

internal language

Or 'I-language'. According to Chomsky, when a child 'learns language', the child attains a state of their individual linguistic capacity – which ultimately consists in the state of their brain – which may differ from child to child. To be distinguished from 'external language' or 'E-language', which is a comparatively vague term for historically conditioned practices of verbalisation such as English, Farsi or ESL.

internal negation

See **external negation**

iterative

See **recursive**

knowledge by acquaintance/knowledge by description

A distinction due to Russell: the former is the direct, conceptually unmediated knowledge of objects and universals; the latter is indirect, conceptually mediated knowledge. Normally, our knowledge of universals (properties and relations) is by acquaintance. Knowledge by description is represented linguistically by means of a definite description involving the concepts via which we think of the object.

knowledge-rule

A candidate for a constitutive rule or norm of assertion: one should assert that p only if one knows that p.

language games

Restriction of **forms of life** to those that essentially involve language, and with an emphasis placed on the 'grammar' of the linguistic activity – on the norms or rules implicit in the activity. Due to Wittgenstein.

locutionary act

In the Austin/Searle scheme, an aspect of a speech-act that is the expression of a meaning. See **illocutionary act**.

logicism

The thesis, made compelling by Frege and Russell, that statements of pure mathematics are really statements of pure logic.

meaning

Classically, once the syntax (the grammar) of a language is settled, meaning is what must be mastered in order to speak the language correctly or appropriately,

and to understand the language. Ordinarily, the term is used in a variety of ways, including its use for 'what one intended', and 'significance' (of some object, event or action). See **sense, reference, implicature** and **content**.

mention

See **use**

metaphysical

A metaphysical statement pertains to the nature or essence of reality and of the things that constitute it, irrespective of point of view. See **epistemological**.

mood

Of a sentence, a matter of the grammar or the order of the words that compose the sentence. The main varieties are the declarative/indicative, the interrogative and the imperative. They are conventionally associated with certain **speech-acts**; normally, one makes a statement by using a sentence in the declarative mood, one asks a question with the interrogative, and one commands using the imperative. See **assertoric force**.

Moore's paradox

Sentences of the form 'p, but I don't believe that p' seem deeply contradictory, but they are not logically contradictory. Unlike 'p and not-p', which cannot be true, the state of affairs that p but one doesn't believe that p is, of course, possible.

multiple relation theory of judgement

Conceived by Russell. The judgement that *The cat is on the mat* is four-termed relation between the judger and the cat, the mat and the relation *is on*. The judgement is a fact. The judgement is true if there is a certain fact not involving the judger, namely the fact that the cat is on the mat. The judgement is false if there is no such fact.

naïve semantics

A simple theory of meaning according to which the meaning of every expression is what it stands for, its referent. At the level of atomic sentences, singular terms mean objects, predicates mean universals (properties and relations) and sentences mean propositions.

naturalism

Basically, a position according to which there are no facts besides those which can be discovered by natural science. However, there is a wide variety of positions which claim to be naturalistic. Perhaps it is relatively common to such positions that naturalism rules out the idea that metaphysics or religion is prior to natural science, that they can be justified in a way which is separate from natural science.

natural kinds

Usually referred to by a noun rather than an adjective, these are kinds or types of individual whose nature, according to Putnam and many others, is determined not by our concepts but in terms of some subset of the properties of the individuals, whether or not the underlying nature is known. There are chemical kinds, biological kinds, geological kinds and so on.

necessary

A necessarily true proposition is one that is true in every possible world; 'necessarily' is equivalent to 'not possibly not'.

negation

Classically, the negation of a sentence is true if the proposition is false, false if it is true. Less classically, the wide-scope negation of a sentence is true if and only if the sentence is not true – leaving open whether the sentence is false or neither-true-nor-false.

negative existentials

The negation of a statement that affirms existence, as in 'Unicorns don't exist'. The negative *singular* existential is a problem: assume for example 'Zeus does not exist' is true; on the face of it, this requires that the predicate '___ does not exist' be true of the referent of 'Zeus'. Therefore there must be such a thing as the referent of 'Zeus', namely Zeus. So Zeus *does* exist. Russell's theory purports to solve this problem; Frege's solves it only artificially.

non-extensional

'Not extensional'; see **extensional**, **intensional** and **hyper-intensional**.

normative, norms

Principles or propositions that are *prescriptive* – that tell one what one ought to do, what one must do, what one is allowed to do or may do. Norms may often be equivalent to *rules*.

observation sentence

An important type of **occasion sentence** where one's disposition to assent to it depends on the stimulation of one's senses *and* where witnesses to the scene tend to agree with the verdict. Examples are 'It's cold' or 'That's a goat'. Relevant primarily to Quine.

occasion sentences

In Davidson and Quine, these are sentences that potentially can change in truth-value, as opposed to *standing sentences*, which cannot. Examples include 'I'm hungry' or 'It's autumn', as opposed to examples such as 'Julius Caesar was assassinated on March 15, 44 BC' or 'Lions are carnivores'.

one-place predicates (monadic predicates)

A **predicate** that has a single empty place which must be filled in to construct a sentence. Examples are '___ drinks wine', 'The first boxer to beat ___ is dead'. If they are explained semantically as having referents, then, depending on the exact theory of meaning we have in view, their referents are properties, concepts, sets (extensions) or functions (from objects to truth-values).

Paradox of Analysis

It is natural to think that a basic activity of analytic philosophy is to find **explicit definitions**. Suppose one advances a definition of X as 'Y and Z'. But a definition

must (by definition!) be such that the two sides of the definition have same content (or sense). No such definition can be informative; 'X = Y and Z' would have the same content as 'X = X'.

parataxis/ctic

Of Davidson's strategy for representing indirect speech and the propositional attitudes, referring to the placing side by side of two clauses – for example, 'Sue said that' and 'Bob smoked' (where 'that' is an indexical, whose reference is the utterance to follow).

Peano Axioms

A characterisation of the natural numbers (NN) advanced by Giuseppe Peano (1858–1932; also by Richard Dedekind, 1831–1916). I. 0 is a NN. II. For every NN, there is a NN that is a successor of it. III. No NN has 0 as a successor. IV. If x and y are NNs, and the successor of x = the successor of y, then x=y. V. If 0 has a certain property, and the successor of every NN has the property, then all NNs have the property. Both Frege and Russell derived Peano's Axioms from their respective definitions of 0 and successor.

perlocutionary act

The effect on the audience, intentional or not, achieved in a speech-act. For example, one performs the **illocutionary act** of asking a question; the perlocutionary act might be to get the audience to answer, or it might be to irritate the audience or prompt them to walk out.

phenomenalism

The stance that only what is directly perceivable is real. All other things are either definable in terms of directly perceived entities or do not exist.

phenomenological reductionism

The thesis that every significant non-analytic statement is in principle reducible to some (possibly complex) statement couched in basic vocabulary that directly describes experiences. Associated with Carnap and with Ayer.

phenomenology

The word is used in a variety of ways, but in analytic philosophy it is used as a *science or discipline or study of experience*. One can speak of the *phenomenology of wine* or the *phenomenology of music*. The characteristic phrase is 'what it is like' – to taste the wine, to hear the music.

picture theory

Due to Wittgenstein in the *Tractatus Logico-Philosophicus*, a theory of meaning for atomic or elementary propositions. Just as a picture of the cat's being on the mat has an element standing for the cat and one standing for the mat, juxtaposed to show the relation, so it is with the proposition, except that the 'standing-for' relation and the showing of the relation are entirely conventional. Propositions, in this usage, are just sentences-in-use.

possible

A proposition that is true in some possible world (including the actual world; is definable as 'not necessarily not'.

pragmatically improper

Utterances that can be semantically correct – which, for example, constitute true statements – but are nonetheless incorrect, because they violate practical or communicative norms. See **cooperative principle**.

pragmatics

A theory of meaning or semantics, in giving the meaning or truth-conditions of sentences, typically leaves many important questions concerning the use of language unanswered; many of these are about the information that can be communicated by the use of sentences but that is not contained in the truth-conditions of sentences. This is standard pragmatics. But some think that pragmatics, as concerning the use of language, takes precedence over semantics or the theory of meaning, or even that semantics can in some sense be reduced to pragmatics.

predicates

An expression with one or more empty places, which becomes a sentence when the empty places are filled with **singular terms**. Examples: ‘__ is a fine boy’; ‘__, Groucho Marx and __ came to stay’. See **one-place predicates**, **two-place predicates**.

predication, ‘is’ of

The English word ‘is’ performs two functions. In its use as the ‘is’ of identity, it joins two singular terms, as when we say ‘George Eliot is Marian Evans’. This use is that of a two-place predicate, satisfactorily rendered as ‘__ = __’, or ‘__ is identical with __’. In its use as the ‘is’ predication, it converts a general term into a one-place predicate. For example, we can convert the general term ‘hungry’ into the predicate ‘is hungry’, which may in turn be conjoined with a singular term, ‘Tina’ for example, to make the sentence ‘Tina is hungry’. The ‘is’ of predication, it can be argued, is dispensable: since we have general terms such as ‘grins’ which do not require the word ‘is’ to make sentence (like ‘Tina grins’), there is no logical reason why we can’t refashion language to make all general terms into terms like ‘grins’. A standard view is that this not the case with the ‘is’ of identity; in this case we have a genuine relation, one that each object bears to itself and to no other.

presupposes

Whereas P *logically entails* Q if and only if it is impossible for P to be true but Q false, P *presupposes* Q if and only if it is impossible for Q to be false and P to be *either* true or false. In other words, the truth of Q is a necessary condition of P’s having any truth-value at all.

principle of charity

Many people including Davidson believe that it is a necessary condition of interpretation or radical interpretation that the subject’s beliefs be mostly true, or at least that their beliefs about obvious things be generally true. Otherwise there is no basis for treating the subject as a rational agent, with beliefs and desires, and no interpretation is possible. Often teamed with a principle of ‘rational accommodation’ – that a person’s propositional attitudes must in the main be coherent.

private language

A language whose referents are necessarily known only to a single person. Wittgenstein famously thought that no such language is possible.

proposition

Yielded by sentence-meaning and a context of utterance. Classically, they are the objects of **propositional attitudes** and fundamental bearers of truth and falsity; their existence and identity are independent of particular languages.

propositional attitudes

Belief, wondering whether, desire, hope and so on may be construed as attitudes towards propositions. Belief, for example, is often construed simply in terms of an ordinary binary (two-place) predicate ‘__ believes __’, indicating a relation between a believer and a proposition.

pure indexicals

See **indexicals**

pure predicate

A **predicate** which contains no **quantifiers**, no **sentence-connectives** and no **singular terms**.

quantifiers

Expressions that indicate how much or how many of a kind of thing there is or are. Examples are ‘each’, ‘every’, ‘some’, ‘a’, ‘many’, ‘most’, ‘much’, ‘seven’ and so on. Attached to a noun or noun phrase – a description in Russell’s vocabulary – we get expressions such as ‘Every man’, which can be inserted into the blank of a predicate to form such sentences as ‘Every man is brave’. This commonality with singular terms once led people on the fruitless search for the referents of such expressions as ‘Every man’, ‘Everything’ and ‘No man’. Frege and Russell recognised quantifiers as a separate category of expression, working together with **variables** (pronouns in ordinary language) and **predicates**.

radical interpretation/radical translation

A procedure where, beginning with no relevant knowledge of a language, one arrives at a translation or interpretation of the language. Unlike Quine’s radical translation, Davidson’s radical interpretation requires the explicit specification of the truth-conditions of each foreign sentence.

recursive, recursion

A recursive process or rule is one that applies to any example of a certain kind of thing, and produces something of the same kind. For example, ‘attaching a bead to a string of beads’ would be a recursive process. For another, ‘+2’ is known as a recursive function on the set of natural numbers because it takes any natural number as input and yields another natural number as output, which can, in turn, be taken as an input for the function, and so on. Linguistic examples include ‘It is not the case that __’ and ‘the mother of __’.

reductionism

See **phenomenological reductionism**

reference

The relation of an expression standing for, or denoting, some item (paradigmatically, an object in the case of singular terms: the relation of ‘Fido’ to Fido).

referentially opaque/referentially transparent

In the sentence 'John believes that the president of the club is dishonest', the position occupied by 'John' is referentially *transparent* – it is open to substitution by **co-referential** terms, whereas the positions that come after 'believes that' are referentially *opaque* – they are not open to substitution by **co-referentials**. See **propositional attitude**, **extensional**.

referential use

Of a definite description, as used by Keith Donnellan: one uses a description *referentially* to refer to an object, if, were the object to fail to satisfy the description, one nevertheless says something about that object. Otherwise the use is *attributive*, in which case Russell's account applies. See **definite description**.

regimentation

Quine's term; to regiment is to devise not strictly an artificial language, but a maximally streamlined language containing only predicates, truth-functions, variables and quantifiers capable of precisely and economically representing scientific theories.

satisfy

A relation between an object and a predicate, the converse of true-of: an object satisfies the predicate if and only if the predicate is true of the object. More technically, it is a relation between an infinite sequence of objects and a sentence, open or closed (Tarski).

scope: wide, narrow

A relation between expressions generally, but especially between logical operators and the parts of sentences which they govern; in 'If Harold was late then Susan is not happy', 'not' takes *narrow* scope (the scope is 'Susan is happy') with respect to 'if-then', whereas 'if-then' takes *wide* scope (its scope is the two sentences – 'Harold is late' and 'Susan is not happy' – out of which the final sentence is composed) with respect to 'not'. In natural language, scope is sometimes ambiguous; for example, 'Everyone did not qualify': one reading is that 'Each person is such that they did not qualify' (the scope of negation inside that of the quantifier); another reading is 'It is not the case that everyone qualified' (the scope of the quantifier inside that of negation).

semantics

The meaning of words or expressions. See **pragmatics**, **syntax**.

sense

In a Fregean environment, every expression has a sense. The sense of a **singular term** such as 'Mars' is a condition that an object must satisfy to be the referent of the term. Alternatively, the sense is a mode of presentation of the referent; it **determines** the referent. The sense of a (one-place) **predicate** determines whether or not an object falls under the predicate (the referent of the predicate is thus itself a rule, in particular a function from objects to truth-values). The sense of a (context-free) sentence is a proposition, what Frege called a thought. Some expressions such as 'Pegasus' or 'The man who walked on the moon in 1894' express senses that fail to determine referents.

sense-data

The immediate sensory content of a perceptual state; Russell tended to think of these as indubitable, the state of having them as incorrigible – in the sense that when you have them, there is no room for doubt that you are having them and that you know which you're having (although one can doubt that they correspond to anything, e.g. that one has apple-shaped sense-data does not itself prove that an apple exists). Russell thought (in 1912) that sense-data (and possibly the self) are strictly speaking the only objects with which we can be acquainted, and for which we can have genuine names – logically proper names. Ordinary proper names function not as genuine names but as disguised or abbreviated **definite descriptions**.

sentence-connectives

An expression which joins sentences together to make a single sentence or which attaches to a single sentence to form a longer sentence. Examples of the first type include 'or', 'and', 'if-then' and 'because'; examples of the first type include 'not', 'necessarily' and 'It is true that' (some of these have other uses, as in forming collective subjects using 'and', such as 'John and Susan'). Some can be plausibly explained as **truth-functional**, but others cannot be; for example, the truth-value of P does not settle the truth-value of 'Necessarily, P'.

singular proposition

See **atomic proposition**

singular term

An expression whose function is to refer to a single object; in the extensional case, filling a one-place **predicate** with a singular term results in a sentence which is true if and only if the predicate is true of the referent of the singular term. Singular terms include proper names such as 'Moscow' and 'Barack Obama', definite descriptions such as 'the present King of England' and indexicals such as 'you' and 'here'.

slurs

Appellations like 'dago' or 'bitch', which seem to refer to a target group but which tend to disparage members of the group, often for just being a member of the group.

speech-act

An act of speaking; there are many kinds of speech-act – telling, promising, questioning, berating, suggesting, joking, requesting and so on, each with a distinctive set of norms and expectations – but the most basic are acts of assertion, command and questioning. See **mood**, **assertoric**, **language games**.

S-structure

In mid-period Chomskyeen Universal Grammar (UG), a level of grammatical representation brought about by grammatical transformations, which provides items to systems of phonological and logical form.

standing sentences

See **occasion sentences**

statement

There are different ways of understanding the word 'statement', but according to one way: relative to a particular context – minimally a particular time and place – a declarative sentence expresses a **proposition** whether or not the sentence is actually uttered in that context; a statement requires the context plus an *actual* utterance of a declarative sentence. See **context of utterance**.

substitutivity

A property of sub-sentential positions or contexts such that any two expressions with the same referent can be exchanged, the one for the other, and the truth-value of the containing sentence will not change. For example, if the **singular terms** *a* and *b* each have the same referent, then the principle of substitutivity holds for the predicate '*F*_' if and only if *Fa* and *Fb* have the same truth-value. See **extensional**.

syntax

A precise description that includes much of what has traditionally been known as grammar. The syntax of a sentence is its particular arrangement or structure of words, described purely formally or morphologically, ideally without reference to semantics or meaning; likewise we can speak of the syntax of the entirety of a language. The syntax of natural language is an empirical theory; the syntax of artificial or formal languages is stipulative. In the philosophy of language, the traditional practice is to consider partially idealised forms of language, with a simplified syntax.

tense

An indexical feature of most ordinary sentences, with the basic forms being *past*, *present* and *future*: 'The dog barked', 'The dog is barking' and 'The dog will bark'. The language of mathematics and some scientific languages or notations, such as that of chemistry, lack distinctions of tense.

thick concept

A concept which expresses both a description and an evaluation; examples are 'courageous', 'wimp', 'scatterbrain'.

truth-condition

A set of all those circumstances under which a given sentence is true.

truth-functional

A sentential connective that results, when attached to one or more sentences, in a sentence whose truth-value is entirely determined by those of the attached sentences. See **extensional**.

T-sentences

Made famous in logic by Tarski, and in the philosophy of language by Davidson. A T-sentence is "'Le neige est blanche" is true if and only if snow is white'. More generally, they are of the form '*s* is true if and only if *p*'. Tarski was able to define truth for a particular language by assuming meaning, in particular the translation of object language to metalanguage, thus that '*s*' refers to a sentence that means

the same as 'p'. Davidson inverts the idea, arguing that we can explain meaning if we presuppose the general concept of truth.

two-place or binary predicate

A **predicate** that has two empty places where **singular terms** go to construct a sentence. Examples are '___kissed___', 'Sally is between ___ and ___'. If they are explained as referring, then, depending on the theory of meaning we assume, they refer to (binary) relations, concepts, sets of ordered pairs, or functions (from pairs of objects to truth-values).

types or tokens

'My cat hates your cat' contains five word-tokens but four words counted as types. Examples involving homonyms such as 'She left on the left' shows that we have to make the type/token distinction among words considered as syntactical types as well as words considered as semantical types.

Universal Grammar

Or 'UG'. The theory, due to Chomsky, of a genetically determined abstract template common to all human languages.

universal quantifier

'All', 'every' etc. See **quantifier** and **existential quantifier**.

universals

Classically objects or individuals have, or bear, properties, and stand in relations to one another. Properties and relations are types of universal. Universals, according to one dominant strand of thought that traces back to Plato, are abstract, necessarily existing, and have no spatial or temporal location. They are what predicates mean, again according to one dominant way of thinking; they are ways that objects can be.

use

(1) We speak of *using* words with their ordinary referents (if they have them) as opposed to *mentioning* them, which is to talk not about ordinary referents but about the words themselves. Paradigmatically, the latter is achieved by using quotation marks (inverted commas). (2) To use language is more broadly to employ its words for some purpose. Some philosophers think that in some sense the meaning of words emerges from, is implicit in or is supervenient on their use; others think that meaning reduces to use, or that for theoretical purposes it is better to speak plainly of use, letting meaning fall where it may. But according to the classical picture, the use of language is sharply distinct from the meaning of its expressions (**pragmatics** should be kept apart from **semantics**). See also **speech-acts**.

variables

In logic and mathematics, one uses 'x' and 'y' and so on for an object that is left unspecified. A *free* variable, in the standard treatment, has a particular *value* – a **referent** – only relative to an assignment of objects from some domain to the variables of the (formal) language (the free variable under an assignment is a kind of

temporary name). A sentence that contains at least one free variable is an *open* sentence. A *closed* sentence is one with no free variables; if it has variables, then those variables are *bound*, normally by quantifiers. For example, if a sentence contains just one free variable, and comes out true under every assignment of values to that variable, then the result of prefixing a universal quantifier governing that variable to it is true (under any assignment); if it comes out false for some assignment then the universal quantification is false. Closely analogous are certain aspects of the behaviour of pronouns in natural language. 'If it is a mammal then it has warm blood' makes a true statement whatever is assigned to 'it'; thus we have the true generalisation 'Everything is such that if it is a mammal then it has warm blood', or more colloquially 'All mammals have warm blood' or more loosely 'Mammals are warm-blooded'.

verification principle

In Ayer's formulation: 'not ... that [a statement] should be conclusively verifiable, but that some possible sense-experience should be relevant to the determination of its truth or falsehood'.

X-phi

An activity of seeking answers to philosophical questions by polling the public, especially questions in the form of thought-experiments. The answers are often interestingly different from the answers given by philosophers, who sometimes say that their answers are the deliverances of 'intuition'. The approach is obviously relevant to questions about the meanings of words, since many people think that the meanings of words are public property.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

• bibliography

- Ahmed, A. (2010) *Wittgenstein's Philosophical Investigations*. London: Continuum.
- Anderson, L. and E. Lepore (2013) 'Slurring Words,' *Nous*, 47(1): 25–48.
- Antony, L. and N. Hornstein, eds (2003) *Chomsky and His Critics*. Oxford: Blackwell.
- Austin, J. (1961) 'Performative Utterances,' in *Philosophical Papers*. Oxford: Clarendon Press.
- Austin, J. L. (1962) *How to Do Things With Words*, Second Edition, J. O. Urmson & M. Sbisá (eds). Cambridge, MA: Harvard University Press.
- Ayer, A. J. (1946) *Language, Truth and Logic*, Revised Edition. London: Victor Gollancz.
- Ayer, A. J. editor (1959) *Logical Positivism*. New York: MacMillan.
- Bach, K. (2018) 'Loaded Words: On the Semantics and Pragmatics of Slurs', in D. Sosa (2018), pp. 60–76.
- Barcan Marcus, R. (1947) 'The Identity of Individuals in a Strict Functional Calculus of Second Order,' *Journal of Symbolic Logic*, 12(1): 12–15.
- Barcan Marcus, R. (1961) 'Modalities and Intensional Languages,' *Synthese*, 13(4): 303–322.
- Barwise J. and J. Etchemendy (1987) *The Liar*. Oxford: Oxford University Press.
- Baz, A. (2012) *When Words are Called For: In Defense of Ordinary Language Philosophy*. Cambridge, MA: Harvard University Press.
- Baz, A. (2017) *The Crisis of Method in Contemporary Analytic Philosophy*. Oxford: Oxford University Press.
- Berto, F. and Jago, M., 2019, *Impossible Worlds*, Oxford: Oxford University Press.
- Berwick, R. & N. Chomsky, 2016. *Why Only us: Language and Evolution*. Cambridge, MA: M.I.T. Press.
- Bezuidenhout, A. (2006) 'The Coherence of Contextualism,' *Mind and Language*, 21(1): 1–10.
- Borg, E. (2012) *Pursuing Meaning*. Oxford: Oxford University Press.
- Brandom, R. (2000) *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.
- Brown, J. (2013) 'Experimental-Philosophy, Contextualism and Subject-Sensitive Invariantism', *Philosophy and Phenomenological Research*, 86(2): 233–494.
- Brown, J. & H. Cappelen, eds (2011) *Assertion: New Philosophical Essays*. Oxford: Oxford University Press.
- Burgess, A., H. Cappelen & D. Plunkett, eds (2020) *Conceptual Engineering and Conceptual Ethics*. Oxford: Oxford University Press.
- Camp, E. (2013) 'Slurring Perspectives,' *Analytic Philosophy* 54(3): 330–349.
- Camp, E. (2018) 'A Dual Act Analysis of Slurs', in D. Sosa ed. 2017, pp.29–59. Oxford University Press
- Cappelen, H. (2018) *Fixing Language: An Essay on Conceptual Engineering*. Oxford: Oxford University Press.

- Cappelen H. and E. Lepore (2005) *Insensitive Semantics: A Defense of Semantic Minimalism and Speech-Act Pluralism*. Oxford: Blackwell.
- Carnap, R. (1936) 'Testability and Meaning'. *Philosophy of Science* 3(4): 419–471.
- Carnap, R. (1937) *The Logical Syntax of Language* [1934]. A. Smeaton, Countess von Zeppelin (tr.). International Library of Psychology, Philosophy and Scientific Method, gen. ed., C. K. Ogden. London: Kegan Paul Trench, Trubner & Co.
- Carnap, R. (1942) *Introduction to Semantics*. Cambridge, MA: Harvard University Press.
- Carnap, R. (1950) 'Empiricism, Semantics and Ontology'. In Carnap (1956).
- Carnap, R. (1956) *Meaning and Necessity: A Study in Semantics and Modal Logic*, Second Edition. Chicago and London: Chicago University Press.
- Carnap, R. (1962) *Logical Foundations of Probability* [1950]. Chicago: University of Chicago Press.
- Carnap, R. (1967) *The Logical Structure of the World* [1928] and *Pseudoproblems in Philosophy*, R. George (tr.). Berkeley: University of California Press.
- Carnap, R. (1995) *The Unity of Science* [1932, 1934] M. Black (tr.). Bristol: Thoemmes Press.
- Cavell, Stanley (1979) *The Claim of Reason* (New York: Oxford University Press).
- Church, A. (1950) 'On Carnap's Analysis of Statement of Assertion and Belief'. *Analysis*, 10(5): 97–99.
- Church, A. (1951) 'A Formulation of the Logic of Sense and Denotation,' in P. Henle (ed.), *Structure, Method and Meaning*. New York: The Liberal Arts Press, pp. 3–24.
- Church, A. (1954) 'Intensional Isomorphism and Identity of Belief,' *Philosophical Studies* 5(5): 65–73.
- Chomsky, N. (1959) 'A Review of B. F. Skinner's *Verbal Behavior*.' *Language* 35(1): 26–58.
- Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: M.I.T. Press.
- Chomsky, N. (1966) *Cartesian Linguistics*. Cambridge, Mass.: M.I.T. Press.
- Chomsky, N. (1968) 'Quine's Empirical Assumptions', *Synthese* 19: 53–68. Reprinted in Davidson and Hintikka 1969.
- Chomsky, N. (1972) *Language and Mind*, enlarged edition. New York, New York: Harcourt, Brace, Jovanivitch.
- Chomsky, N. (1975) *Reflections on Language*. New York, New York: Fontana, Random House.
- Chomsky, N. (1980) *Rules and Representations*. Oxford: Blackwell.
- Chomsky, N. (1986) *Knowledge of Language: Its Nature, Origin and Use*. New York, New York: Praeger.
- Chomsky, N. (1995) *The Minimalist Program*. Cambridge, MA: M.I.T. Press.
- Chomsky, N. (2000) *New Horizons in the Study of Language*. Cambridge: Cambridge University Press
- Coffa, A. (1991) *The Semantic Tradition from Kant to Carnap: To the Vienna Station*, L. Wessels (ed.). Cambridge: Cambridge University Press.
- Davidson, D. (1984) *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
- Davidson, D. (2006) *The Essential Davidson*, K. Ludwig & E. Lepore (eds). Oxford: Oxford University Press.
- Davidson, D. & G. Harman (1972) *Semantics of Natural Language*. Dordrecht: Reidel.
- DeRose, K. (2002) 'Assertion, Knowledge and Context,' *The Philosophical Review*, 111(2): 167–203.
- DeRose, K. (2011) 'Contextualism, Contrastivism, and X-Phi Surveys', *Philosophical Studies*, 156(1): 81–110.
- Deutsch, M. (2009) 'Experimental Philosophy and the Theory of Reference', *Mind and Language*, 24(4): 445–466.
- Deutsch, M. (2015) *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method*. Cambridge, MA: M.I.T. Press.

- Diaz-Leon, E. (2019) 'Pejorative Terms and the Semantic Strategy', in *Acta Analytica* 35(1): 23–34.
- Diamond, C. (1988) 'Losing Your Concepts', *Ethics*, 98(2): 255–277.
- Donnellan, K. (1966) 'Reference and Definite Descriptions,' *Philosophical Review*, 75(3): 281–304.
- Dummett, M. (1973) *Frege: Philosophy of Language*. London: Duckworth.
- Dummett, M. (1978) *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press.
- Dummett, M. (1993a) *Frege: Philosophy of Language*, Second Edition. Cambridge, MA: Harvard University Press.
- Dummett, M. (1993b) *The Seas of Language*. Oxford: Clarendon.
- Eklund, M. (2020) 'Conceptual Engineering in Philosophy', in Burgess et. al.: 15–30.
- Evans, G. (1982) *The Varieties of Reference*, J. McDowell (ed.), Oxford: Clarendon Press.
- Evans, G. (1985) *Collected Papers*. Oxford: Oxford University Press.
- Everett, A. & T. Hofweber, eds (2000) *Empty Names, Fiction and the Puzzles of Non-Existence*. Stanford: CSLI Publications.
- Everett, D. L. (2008) *Don't Sleep, There are Snakes*. New York: Pantheon Books.
- Evnine, S. (1991) *Donald Davidson*. Cambridge: Polity Press.
- Feltz, A. and E. T. Cokely (2009) 'Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism.' *Consciousness and Cognition* 18(1): 342–350.
- Fine, K. (1994) 'Essence and Modality'. *Philosophical Perspectives*, 8: 1–16.
- Fodor, J. (1987) *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.
- Frege, G (1984) 'Compound Thoughts' [1923–26], in B. McGuinness (ed.), M. Black et al (trs), *Collected Papers on Mathematics, Logic, and Philosophy*. Oxford: Basil Blackwell.
- Frege, G. (1974) *Foundations of Arithmetic* [1884] J. L. Austin (tr.). Oxford: Basil Blackwell.
- Frege, G. (1979) *Posthumous Writings*, H. Hermes et al (eds and trs). Oxford: Basil Blackwell.
- Frege, G. (1997) *The Frege Reader*, M. Beaney (ed.). Oxford: Wiley-Blackwell.
- Fricker, M. (2007) *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.
- García-Carpintero, M. (2017) 'The philosophical significance of the De Se.' *Inquiry*, 60(3): 253–276.
- Geach, P. (1967) 'Intentional Identity,' *Journal of Philosophy*, 64(20): 627–632.
- Goldberg, S., ed (2020) *The Oxford Handbook of Assertion*. Oxford: Oxford University Press.
- Grice, H. P. (1989) *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Grice, H. P. & P. F. Strawson (1956) 'In Defense of a Dogma,' *The Philosophical Review*, 65(2): 141–158.
- Häggqvist S. & Å. Wikforss (2018) 'Natural Kinds and Natural Kind Terms: Myth and Reality', *British Journal for the Philosophy of Science*, 69(4): 911–933.
- Harris, D., D. Fogal & M. Moss editors (2018). *New Work on Speech Acts*. Oxford: Oxford University Press.
- Haslanger, S. (2000) 'Gender, Race: (What) Are They? (What) Do We Want Them To Be?' *Noûs*, 34(1): 31–55.
- Hawthorne, J. & D. Manley (2012) *The Reference Book*. Oxford: Oxford University Press.
- Hedger, J. (2012) 'The semantics of racial slurs: Using Kaplan's framework to provide a theory of the meaning of derogatory epithets'. *Linguistic and Philosophical Investigations*, 11: 74–84.
- Hills, D. (2016) 'Metaphor,' in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (<https://plato.stanford.edu/entries/metaphor>).

- Hom, C. (2008) 'The Semantics of Racial Epithets,' *Journal of Philosophy*, 105(8): 416–440.
- Horne, Z. & J. Livengood (2017) 'Ordering Effects, Updating Effects, and the Specter of Global Skepticism,' *Synthese*, 194(4): 1189–1218.
- Hornsby J. (1993) 'Speech acts and pornography,' *Women's PhilosRev*, 10: 38–45.
- Husserl, E. (1973) *Logical Investigations* [1900/1; revised 1913] J. N. Findlay (tr), London: Routledge.
- Hylton, P. & G. Kemp 'Willard van Orman Quine,' in E. N. Zalta (ed), *Stanford Encyclopedia of Philosophy* (<http://plato.stanford.edu/entries/quine>).
- Hylton, P. (2005) *Propositions, Functions, and Analysis: Selected Essays on Russell's Philosophy*. Oxford: Oxford University Press.
- Jenkins, K. (2016) 'Amelioration and Inclusion: Gender Identity and the Concept of Woman,' *Ethics*, 126(2): 394–421.
- Jylkkä, J., H. Railo & J. Haukioja (2009) 'Psychological Essentialism and Semantic Externalism: Evidence for Externalism in Lay Speakers' Language Use', *Philosophical Psychology*, 22(1): 37–60.
- Kaplan, D. (1969) 'Quantifying In,' in D. Davidson & G. Harman (eds), *Word and Objections: Essays on the Work of W. V. Quine*. Dordrecht: Reidel.
- Kaplan, D. (1989) 'Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals,' in J. Almog, J. Perry & H. Wettstein (eds), *Themes from Kaplan*. Oxford: Oxford University Press.
- Kaplan, D. (2004) 'The Meaning of Ouch and Oops' (unpublished transcription of the Howison Lecture in Philosophy at U.C. Berkeley, transcribed by E. Coppock).
- Katz, J. (1981) *Language and Other Abstract Objects*. Oxford: Basil Blackwell.
- Kemp, G. (2023) *Quine's Philosophy: An Introduction*. London: Bloomsbury.
- Khoo, J. & R. K. Sterke, eds (2021) *The Routledge Handbook of Social and Political Philosophy of Language*. London: Routledge.
- Kim, M. & Y. Yuan (2015) 'No Cross-Cultural Differences in the Gettier Car Case Intuition: a Replication Study of Weinberg et al. 2001', *Episteme*, 12(3): 355–361.
- Knobe, J. & S. Nichols (2008) *Experimental Philosophy*. Oxford: Oxford University Press.
- Knobe, J. & S. Nichols (2017) 'Experimental Philosophy', in the *Stanford Encyclopedia of Philosophy* (<https://plato.stanford.edu/entries/experimental-philosophy>).
- Kripke, S. (1971) 'Identity and Necessity,' in M. Munitz (ed), *Identity and Individuation*. New York: New York University Press, pp. 135–164.
- Kripke, S. (1977) 'Speaker's Reference and Semantic Reference,' in P. French, T. E. Uehling Jr & H. K. Wettstein (eds), *Studies in the Philosophy of Language*. Bloomington: University of Minnesota Press, pp. 255–276.
- Kripke, S. (1979) 'A Puzzle About Belief,' in A. Margalit (ed), *Meaning in Use*. Dordrecht: Reidel, pp. 239–283.
- Kripke, S. (1980) *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Kripke, S. (1982) *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press.
- Kripke, S. (2013) *Reference and Existence*. Oxford: Oxford University Press.
- Langton R. (1993) Speech acts and unspeakable acts. *Philosophy and Public Affairs* 22(4): 292–330
- Larson, R. K. & P. Ludlow (1993) 'Interpreted Logical Forms', *Synthese*, 95(3): 305–355.
- Lepore, E. & M. Stone (2014) *Imagination and Convention*. Oxford: Blackwell.
- Lepore, E. & U. Stojnić (2022) 'Inescapable articulations: Vessels of lexical effects'. *Nous*, 56(3): 742–760.
- Lewis, D. (1983) *Philosophical Papers*, Vol. I. Oxford: Oxford University Press.
- Lewis, D. (1998) *Papers in Philosophical Logic*. Cambridge: Cambridge University Press.

- Linsky, L. (1971) *Reference and Modality*. London: Oxford University Press.
- Locke, J. (1961) *An Essay Concerning Human Understanding*. London: Everyman's.
- Ludlow, P. (2013) *The Philosophy of Generative Linguistics* (Oxford: Oxford University Press).
- Lugg, A. (2000) *Wittgenstein's Investigations 1–133: A Guide and Interpretation*. London: Routledge.
- MacFarlane, J. (2014) *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford: Oxford University Press.
- Machery, E., R. Mallon, S. Nichols & S. Stich (2004) 'Semantics, Cross-Cultural Style', *Cognition*, 92(3): 1–12.
- Machery, E., S. Stich, D. Rose, A. Chatterjee, K. Karasawa, N. Struchiner, S. Sirker, N. Usui & T. Hashimoto (2015) 'Gettier Across Cultures', *Noûs*, 51(3): 645–664.
- MacKinnon C. (1987) *Feminism unmodified. Discourses on life and law*. Harvard University Press, Cambridge.
- Malpas, J. 'Donald Davidson,' in E. N. Zalta (ed), *Stanford Encyclopedia of Philosophy* (Fall 2010 Edition). <http://plato.stanford.edu/entries/davidson>
- Mates, B. (1952) 'Synonymity,' in L. Linsky (ed), *Semantics and the Philosophy of Language*. Urbana, IL: University of Illinois Press, pp. 111–136.
- Meinong, A. (1981) 'The Theory of Objects' in R. Chisholm (ed), *Realism and the Background of Phenomenology*. Atascadero, CA: Ridgeview, pp. 76–117.
- Mill, J. S. (1963) [1843] 'A System of Logic, Ratiocinative and Inductive,' in J. M. Robson (ed), *Collected Works of John Stuart Mill*, Vol. 7–8. Toronto: University of Toronto Press.
- Montague, R. (1974) *Formal Philosophy. Selected Papers of Richard Montague*, R. Thomason (ed). New Haven: Yale University Press.
- Myers-Schulz, B. & E. Schwitzgebel (2013) 'Knowing that P Without Believing that P', *Noûs*, 47(2): 371–384.
- Neale, S. (1993) *Descriptions*. Cambridge, MA: MIT Press.
- Neurath, O. (1959) 'Protocol Sentences' [1932], G. Schick (tr), in A. Ayer (ed), *Logical Positivism*. New York: MacMillan, pp. 199–208.
- Nietzsche, F. (1968) *The Will to Power* [1901] W. Kaufmann (tr), New York: Random House.
- Noonan, H. (2001) *Frege: A Critical Introduction*. Cambridge: Polity Press.
- Parsons, T. (1975) 'A Meinongian Analysis of Fictional Objects,' *Grazer Philosophische Studien*, 1: 73–86.
- Parsons, T. (1980) *Nonexistent Objects*. New Haven: Yale University Press.
- Peregrin, J. (2013) *Inferentialism: Why Rules Matter*. New York: Palgrave.
- Perry, J. (1977) 'Frege on Demonstratives,' *Philosophical Review*, 86(4): 474–497.
- Perry, J. (1979) 'The Problem of the Essential Indexical,' *Noûs*, 13: 3–21.
- Pierce C. S. (1934) 'Belief and Judgment' [1877], in C. Hartshorne & P. Weiss (eds), *Collected Papers*, Vol. V. Cambridge, MA: Harvard.
- Pope-Wyatt, M. 'Slurs, Pejoratives, and Hate Speech,' *Oxford Bibliographies On-Line* (<https://www.oxfordbibliographies.com/display/document/obo-9780195396577/obo-9780195396577-0403.xml>).
- Priest, G. (2005) *Towards Non-Being: The Logic and Metaphysics of Intentionality*. Oxford: Clarendon.
- Putnam, H. (1975) 'The Meaning of "Meaning"', *Minnesota Studies in the Philosophy of Science*, 7: 131–193.
- Quine, W. V. (1960) *Word and Object*. Cambridge, MA: M.I.T. Press.
- Quine, W. V. (1961) 'Two Dogmas of Empiricism,' in *From a Logical Point of View*, Second Edition. Cambridge, MA: Harvard University Press, pp. 20–46.
- Quine, W. V. (1969) 'Speaking of Objects,' in *Ontological Relativity and Other Essays*. New York and London: Columbia University Press, pp. 1–25.

- Quine, W. V. (2008) 'Reasons for the Indeterminacy of Translation' [1970], in *Confession of a Confirmed Extensionalist* Cambridge, MA: Harvard University Press, pp. 209–214.
- Quine, W. V. (1974) *The Roots of Reference*. La Salle, IL: Open Court.
- Quine, W. V. (1975) 'Truth by Convention' [1935], in *The Ways of Paradox*, Revised Edition. Cambridge, MA: Harvard University Press, pp. 77–106.
- Quine, W. V. (1975) 'Quantifiers and Propositional Attitudes,' in *Ways of Paradox*, Revised Edition. Cambridge, MA: Harvard University Press, pp. 185–196.
- Quine, W. V. (1981) 'Things and Their Place in Theories,' in *Theories and Things*. Cambridge, MA: Harvard University Press, pp. 1–23.
- Quine, W. V. (1992) *Pursuit of Truth*, Revised Edition. Cambridge, MA: Harvard University Press.
- Rey, G. (2020) *Representation of Language: Philosophical Issues in a Chomskyan Linguistics*. Oxford: Oxford University Press.
- Richardson, A. & T. Uebel, eds (2007) *The Cambridge Companion to Logical Empiricism*. Series: Cambridge Companions to Philosophy. Cambridge: Cambridge University Press.
- Russell, B. (1903) *Principles of Mathematics*. Cambridge: Cambridge University Press.
- Russell, B. (1905) 'On Denoting,' *Mind*, 14(56): 479–493.
- Russell, B. (1911) 'Knowledge by Acquaintance and Knowledge by Description,' *Proceedings of the Aristotelian Society, New Series*, 11: 108–128.
- Russell, B. (1918) *Mysticism and Logic and Other Essays*. London: George Allen and Unwin.
- Russell, B. (1921) *The Analysis of Mind*. London: George Allen and Unwin.
- Russell, B. (1927) *The Analysis of Matter*. London: Routledge and Kegan Paul.
- Russell, B. (1940) *Inquiry into Meaning and Truth*. New York: Norton.
- Russell, B. (1948) *Human knowledge, its scope and limits*. New York: Simon & Schuster.
- Russell, B. (1956) *Logic and Knowledge*. London: George Allen and Unwin.
- Russell, B. (1973) *Essays in Analysis*. London: George Allen and Unwin.
- Russell, B. (1985) *The Philosophy of Logical Atomism* [1918, 1924], D. Pears (ed). La Salle, IL: Open Court.
- Sainsbury, M. (1979) *Russell*. London: Routledge.
- Salmon, N. (1986, 1991) *Frege's Puzzle*. Atascadero, CA: Ridgeview.
- Salmon, N. (1989) 'How to Become a Millian Heir,' *Noûs*, 23(2): 211–220.
- Schlick, M. (1959) 'The Foundation of Knowledge' [1932], D. Rynin (tr), in A. Ayer (ed), *Logical Positivism*. New York: MacMillan, pp. 209–227.
- Schwitzgebel, E. & F. Cushman (2015) 'Philosophers' Biased Judgments Persist Despite Training, Expertise and Reflection', *Cognition*, 141: 127–137.
- Searle, J. (1969) *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Searle, J. (1979) 'Metaphor,' in *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
- Sellars W. (1956) 'Empiricism and the Philosophy of Mind,' in H. Feigl & M. Scriven (eds), *Minnesota Studies in the Philosophy of Science*, Vol. I. Minneapolis, MN: University of Minnesota Press, pp. 253–329.
- Skyrms, B. (1978) 'An Immaculate Conception of Modality', *Journal of Philosophy*, 75: 77–96.
- Sosa, D., ed (2018) *Bad Words: Philosophical Perspectives on Slurs*. Oxford: Oxford University Press.
- Sperber, D. & D. Wilson (1995) *Relevance: Communication and Cognition*, Second Edition. Cambridge, MA: Harvard University Press.
- Stanley, J. (2017) Hale, B., Wright, C., & Miller, A. (eds.) *A Companion to the Philosophy of Language*, second edition. (Oxford: Blackwell), pp. 920–947.
- Strawson, P. F. (1950) 'On Referring,' *Mind*, 59(235): 320–344.

- Strawson, P. F. (1963) 'Carnap's views on constructed systems versus natural languages in analytic philosophy', in P. A. Schilpp (ed) *The Philosophy of Rudolf Carnap* (La Salle: Open Court), pp. 503–518.
- Tarski, A. (1993) 'Truth and Proof' [1969], in R. Hughes (ed), *A Philosophical Companion to First-Order Logic*. Indianapolis: Hackett.
- Travis, C. (2008) *Occasion-Sensitivity: Selected Essays*. Oxford: Oxford University Press.
- Turing, A. M. (1950) 'Computing Machinery and Intelligence'. *Mind*, 49: 433–460.
- Uebel, T. (2007) *Empiricism at the Crossroads. The Vienna Circle's Protocol Sentence Debate*. Chicago: Open Court.
- Whitehead, A. W. and B. Russell (1962) *Principia Mathematica* to *56. Cambridge: Cambridge University Press.
- Whiting, D. (2013) 'It's Not What You Said, It's the Way You Said It: Slurs and Conventional Implicature,' *Analytic Philosophy*, 54(3): 364–377.
- Williamson, T. (1996) 'Knowing and Asserting,' *The Philosophical Review*, 105(4): 489–523.
- Williamson, T. (2000) *Knowledge and Its Limits*. Oxford: Oxford University Press.
- Williamson, T. (2007) *The Philosophy of Philosophy*. Oxford: Blackwell Publishing.
- Wittgenstein, L. (1958) *The Blue and Brown Books*. Oxford: Blackwell Publishers.
- Wittgenstein, L. (1961) *Tractatus Logico-Philosophicus* [1921] D. F. Pears & B. F. McGuinness (trs). New York: Humanities Press.
- Wittgenstein, L. (2009) *Philosophical Investigations* [1953], Fourth Edition P. M. S. Hacker & J. Schulte (eds and trs). Oxford: Wiley-Blackwell.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

• index

Note: Entries in **bold** appear in the Glossary. Page numbers followed by “n” denote endnotes.

- a priori* 1–2, 82, 85, 89, 90, 97,
105, 215; in Ayer 77, 78, 82; in
descriptivist theory of proper
names 124–5; in Frege 31, 48; and
natural kind terms 142, 225; in
Wittgenstein 75
- abstract entities 11, 46–9 *passim*, 70–1,
72, 73, 89, 116, 133, 233, 244
- acquaintance 61–4, 71, 72–3, 102,
103, 115, 132; *see also* knowledge by
acquaintance
- actual world 122–5, 127, 132–4, 144–7,
154, 238, 241
- adverbs 143, 187–8
- analogical substitution 109
- analysis, paradox of 50, 261, 271
- analyticity 6–7, 101; in Ayer 78–82,
90; in Carnap 83, 84, 85; in
Kripke 125–6, 134, 135; in Quine
111–14, 119, 120; in Schlick 87; in
two-dimensional semantics 147
- anaphoric** 140–1, 222–3
- ‘and’ 23, 75–6, 245–6
- anomalist monism 190
- Anscombe, Elizabeth 235
- antecedents 124, 164–5, 252
- applying to things 19
- argument from analogy 80
- argument from error 126
- argument from ignorance 126
- Aristotle 26, 68, 259
- arithmetic 47, 69, 113–14; *see also*
mathematics
- artefacts 143, 265
- assent, disposition to 107–9, 183, 184,
185, 206, 207, 208
- assertion** 120–6, 185, 234–7, 246
- assertoric force 8, 157–8, 235; *see also*
assertion
- atomic propositions** 21–2, 63, 70, 75
- atomic sentences** 16, 18–19, 20–1, 23,
27, 32, 33–5, 61–2, 70, 93, 180–2
- atomistic 285
- attributive use 167, 172
- Austen, Jane 244, 271
- Austin, John L. xvi, 4, 157, 158–61,
167, 171–2, 226, 234, 241, 247, 259,
265, 271
- author of *Waverley* 53, 59–60, 66
- Ayer, Alfred J. xv, 3, 65, 77–82, 83, 84,
85, 87, 89, 90, 120, 134, 135
- Barcan Marcus, Ruth 27, 127, 135,
137n1
- Baz, Avner 242, 271
- behaviourism 80, 102, 107, 213, 226–7
- belief 11, 22, 199–208; and assertion
235–6, 246; Davidson on 185–7,
188–9; extensionality and
substitutivity in 194–7; Frege on 31,
40–1, 45; Quine on 199–201; *see also*
propositional attitudes
- Bergmann, Gustav 88
- Berkeley, Bishop 26, 63, 78, 79–80
- biconditionality** 24
- binary predicates** *see* two-place
predicates
- Bloomfield, Leonard 216

- Bouwmsma, O. K. 171
 Brandom, Robert 245
 Burge, Tyler xv, xvi, 153, 190
- cancellability 161, 163, 173, 246, 253, 272
 Carnap, Rudolf xv, 3, 65, 77, 78, 82–5, 86, 87, 88, 89, 90, 91, 134–5, 208, 262–4, 272; and Frege 46–7; and Quine 112, 119–20
 Carroll, Lewis 136
 Castañeda, Hector-Neri 153
 causal theory of reference (causal chains) 129, 131, 135–6
 Cavell, Stanley 102, 242, 271
character 146, 153–4
 charity *see* principle of charity
 Charles, Ray 5
 Chomsky, Noam 13, 28n2, 77, 167, 212–32
 Church, Alonzo xv, 44, 134, 208
 circularity 82, 261
circumstance of evaluation 146–8, 154
 cities 97
 closed sentences 25, 182–3
 co-extensive 38–9, 60, 133, 193
cognitive content 3, 130, 132
cognitive meaning 7–8, 31, 83, 85, 120, 264–5
cognitive value 29–31, 36, 48, 52, 53, 126
 colour 63, 76, 94, 248
 colouring 253, 265, 272; *see also* tone
 complex singular terms 16, 31, 54–5, 66
compositionality 12–14, 21, 109;
 Chomsky on 215, 229; Davidson on 176–7; Frege on 13–14, 30, 34–5, 38, 46–9
 compound *see* indexicals
concepts 150–1, 259–68, 259; Frege on 44; Frege's horse problem 43–5
 concrete 124, 243–4
conditionality 23–4, 164–5
conditionals *see* counterfactual conditionals
conjunctions 23, 75–6
connotation 27, 127, 204, 258
 constitutive rules 172, 235–7, 246
 consumers 129
content 3, 5, 9, 41; and belief 31; in Kaplan 146, 154; Kripke on 167–8; as meaning 8, 41, 157–9; of proper names 62, 206–8; sense-contents 79–81; slurs 251–8; *see also* propositions
context of utterance 9–12, 139, 139–54, 158, 165, 166, 168, 171, 194–200, 204, 225, 255
 context principle 45–6
context-relativity, context-sensitivity 11–12, 140, 237–42, 247, 255, 256
contextual definition 56, 80, 90, 91, 125, 261, 264
 contextual indices 140, 145
 contextualism 140–1, 247
contingent 122, 125–6, 130–1, 134, 135, 142, 147, 198, 244, 247, 267
 contradictions 7, 38, 68–9, 87, 115, 174n1, 202, 205
conventional implicature 173–4n1, 275n1
 conventions: of assertion 234–7;
 linguistic rules as 147; logic as 76, 84;
 of mood and force 158–9; slurs 254
conversational implicature 162–6, 168, 171, 172, 186, 253, 254, 272; *see also* conventional implicature
cooperative principle 163–5, 168, 170, 172, 236
 copulation 71
co-referentials 31, 60, 193, 200–1
counterfactual conditionals 124, 127, 149, 174, 240–1, 247
 craving for generality 95
 Creath, Richard 259
 creationism 244
criterion 20, 34, 43, 79, 90, 101–2, 142
 cross-reference devices 140
- Darcy, Mr 244
 Davidson, Donald xvi, 4, 119, 150, 175–92, 208–9, 212
de dicto *see de re*
de re 198–200, 201–2, 206–8, 209
de se belief 203–4, 210
declarative mood 2, 8–9

- deep structure, D-structure 220–1
definite descriptions 54–61, 66,
 80; fixing the reference 130–1;
 intensional semantics 132–4;
 objections to Russell's theory of
 124–7, 131; rigid designation 127–9
 deictic expressions *see* indexicals
demonstration 10, 140, 153, 205, 210
demonstratives 10, 87, 110, 139–40,
 153, 189–90, 205, 209; implicit
 indexicals 154, 204–5; indexicals
 139–41, 152–3
 Dennett, Daniel 119
denotation 56, 63, 64, 115, 127,
 142, 242
 Descartes, Rene 147, 214, 215, 229
 descriptivist theory of proper names
 124–7, 132
determines 32–3, 34–5
direct definition 56, 261; *see also*
 explicit definition
 direct reference theory 127–31
directly referring expressions 62–4,
 127–31, 206–8
 disguised definite descriptions 60–1, 66
disjunctions 23, 75–6, 96, 165, 174n3
 dispositions: to assent 107, 108, 110,
 111, 112, 185, 206; in Wittgenstein
 99–100, 102
 Dodgson, Charles 136
 Donnellan, Keith 167, 172
 dubbing 129, 131, 136
 Duhem, Pierre 79
 Dummett, Michael xv, 47, 50, 51, 78,
 234–5, 245

 Einstein, Albert 88, 229
 Eklund, Matti 268
 elementary propositions *see* atomic
 propositions
 elementary sentences *see* atomic
 sentences
 Eliot, George 60–1, 62, 189
 emotivism 81–2
 empirical 64, 77–80, 82–5, 86, 90,
 106–7, 114, 134, 142–3, 214–15;
 Davidson on 176, 183–5; *see also*
 logical empiricism
 empiricism 77–91, 214–15
empty singular terms 31–2, 36–7, 48,
 242–5
 epistemic contextualism 239–40, 247
 epistemic holism 113–14
 epistemic relativism *see* epistemic
 contextualism
epistemological, epistemology xvi, 8,
 29, 77–91, 105, 113–14, 119, 123,
 126, 132, 134, 148, 200, 201, 271
 epistemological reductionism 77–82,
 83, 113
 equivalence 67n4, 111, 118, 134
 error, argument from 126
 essence 74, 141–3, 156n1, 198–9, 270
 essential indexicals 150–1, 153, 203–4
 Evans, Gareth 131, 153
 Evening Star 29–31, 36, 47–8, 53, 206
 Everett, Daniel 14
 exclusive disjunction *see* disjunctions
 existence problems 37, 52, 57–9, 62;
 see also nonexistence
existential quantifiers 25, 58, 65,
 200, 204
explicit definition 80, 261, 264, 265
explication 114–16, 260, 262–4, 268,
 272
 exportation 199, 201–3, 209
expressive meaning 7–8, 253–4,
 265, 272
extensional 43–4, 178–80; *see also*
 extensionality
 extensionality 38–41, 136–41, 193–4;
 see also substitutivity
extensions 20, 34–6, 43, 47, 48, 60,
 68, 133–4, 154, 156, 187, 193, 252,
 255, 258
external negation 56–7
 external questions 85, 90
 externalism, semantic 143–4, 153,
 257–8

 facts 69–75, 78, 89–90, 212
family resemblance 96–8, 103, 265
 Feigl, Herbert 85, 88
 fictional objects xvi, 4, 242–5, 247
 fictional operators 244
 first-level predicates 58

- first-person pronouns 139, 150–1, 203–4, 209–10
fixing the reference 129–31, 136
 Fodor, Jerry xv, 208
force 8–9, 157–8, 160, 161, 163, 171, 235, 258
forms of life 94–5, 98, 100, 102, 228, 235
 Frank, Philipp 82, 85, 88
 Frege, Gottlob xv, xvi, 3, 4, 13–14, 15, 26, 27, 29–49, 234; on indexicals 151–2, 204–5; on propositional attitudes 194–7, 208; Russell and 65, 68–9, 124–5, 132, 242; on slurs 253–4
 Frege-Geach problem 82
 functions 43–4, 50, 167; *see also* truth-functional

 Galileo 188–9, 216
 games 96, 265, 193; *see also* language-games
 ‘Gavagai’ 108–10, 121, 183–6, 191
 Geach, Peter xv, 82, 90, 135
general terms 19, 143, 270
 generalisations 24–6, 62, 74, 238
 generality 24, 95
 Gettier, Edmund 236, 261, 271
 God, existence of 38
 Gödel, Kurt 6, 88
 Gorbachev, Mikhail 36
 grammar xvi, 2, 9, 16, 18, 45, 77, 81, 97, 187; universal 4, 28n2, 212–32; *see also* syntax
 Grice, Herbert Paul xv, xvi, 4, 157, 171; on assertion 192, 193–4; on the cooperative principle 163–4, 165; on implicature 162–4; on metaphor 169–70; and Quine 112
 guises 206–7, 210

 Hahn, Hans 82, 85
 Harman, Gilbert 199, 135
 Harris, Zelig 216
 Haslanger, Sally 260, 266–8
 Heidegger, Martin 83
 Hempel, Carl 82, 85, 88

 Hilbert, David 46
historical chains 129, 131, 136, 143, 244, 255
 holding true 185–6
holism 65, 76, 219; Davidson on 180, 184, 188; Quine on 113–14, 120
 Homer 131, 243, 244
 house 89, 226
 Humboldt, Alexander von 13, 215
 Hume, David 26, 64, 78, 79, 112, 120, 134, 142, 214, 259
 Husserl, Edmund 46, 50
 Hylton, Peter 119
hyper-intensional 195–6, 197–9; Davidson on 179, 189; fictional operators 244; multiple hyper-intensional embedding 197–8, 209
 hypotheses 77–81

 ideas 19–20, 41
 identity 52–3, 58; ‘is’ 4
 idiolects 205, 208
 iff 24
 ignorance, argument from 126
illocutionary act 160–2, 169–70, 171, 172, 254, 259, 264, 265
 imperative mood 8, 157–8
implicature 162–6, 168, 172, 173–4n1, 186, 253, 254, 272, 133–4, 253
 implicit indexicals 140, 141, 204–5
implicit relativity 237–42, 247
 inclusive disjunction *see* disjunctions
 indeterminacy 243–4; of interpretation 185; of translation 109–11, 117–18
indexicals xvii, 4, 10, 137n4, 139–43, 150–2, 153–4, 156n2&3, 187–9, 190, 203–4, 205, 209–10, 223; context-relativity 237–8, 250; explicit and implicit 204–5
indicative mood *see* declarative mood
 indirect sense and reference 41, 197–8
 Induction Principle 79, 100, 214
inferentialism 234, 245–6, 247–8
 information value *see* cognitive value
 inscrutability of reference 210–11
intensional context 198–9

- intensional isomorphism 134, 208
- intensional operators 123
- intensional semantics 132–4, 136, 152–3
- intention 162
- inter-definability 123
- interdependence of belief and meaning 185–6
- internal negation 57
- internal questions 85, 90
- interpretation 98–100, 183–7
- interrogative mood 8, 157–9, 160–1, 171
- intuitions 89, 165, 225, 268–70, 271, 273
- iterative** *see* recursion

- Jenkins, Katherine 267–8
- jokes 165
- judgement, theory of 71–3, 89
- justified-true-belief 236, 261, 268–9

- Kant, Immanuel 32, 38, 112, 120, 142
- Kaplan, David xv, 4, 127, 134, 135, 137n6, 138n4, 116, 149, 212; on slurs 253–5, 258, 272
- on two-dimensional semantics 146; on ‘vivid designators’ 202–3, 208, 209
- Kilimanjaro, Mt. 175
- Klimt, Gustav 85
- knowledge 236, 239–40, 247, 268–9
- knowledge by acquaintance** 61–4, 66
- knowledge by description** 61–4, 66
- knowledge-rule** 237
- Kripke, Saul xvi, 4, 125–32, 134–6, 141, 143, 146, 148, 168, 172, 204, 208, 212, 244, 259, 270; and Russell 168; *Wittgenstein on Rules and Private Language* 98, 99

- language games** 93–5, 96, 97, 100, 103
- Language of Thought 208
- Le Verrier, Urbain 31, 40, 45, 196–7, 200, 209
- Leibniz, Gottfried Wilhelm 26, 215, 229
- Lepore, Ernie 190
- Lewis, David xv, 119, 124, 134, 136, 153, 204, 241, 247, 269, 273
- linguistic meaning 140, 151, 154, 224–6
- linguistics 2–3, 4, 212–16, 226–8
- Linus (van Pelt) 247
- Loch Ness Monster 32, 48, 53, 57–9
- Locke, John 3, 19, 26, 41, 42, 64, 143
- locutionary acts** 160–1, 172
- logic, Frege’s impact on 46–7
- logical atomism 65, 68–9, 85–6
- logical constructions (Ayer) 79–80
- logical empiricism *see* Logical Positivism
- logical form 55–6, 65, 76, 90
- Logical Positivism (or logical empiricism) 3, 4, 69, 77–88, 89, 90–1
- logical syntax and operators 21–6
- logically proper names 56, 61–2, 63, 66, 76, 132, 225
- logicism** 68–9
- Loos, Alfred 85
- ‘L-state’ 152
- lotteries 268–9
- Lycan, William xvi

- Mach, Ernst 78, 85
- Madagascar 131, 138n7
- malapropisms 190
- Malcolm, Norman 171
- marriage 161, 235, 268
- material objects 65, 78, 79–80
- mathematics xvi, 2, 3, 26, 27; as logic 46–7, 65, 68–9; as necessary truths 122, 198; Quine on 113–14, 115
- Maurusius, Emperor 11–12
- maxims of quality, manner, etc. 163, 165, 168, 171, 172, 186, 236
- McDowell, John 91, 102, 190
- meaning** 2–4, 7–13, 157; Chomsky on 224–6; Davidson on 175–91; Frege on 33; naïve semantics of 15–22; Quine on 105–6, 109–12, 117–18; in slurs 251–2, 253–5; Wittgenstein on 90, 91–5, 101–2, 103
- Meinong, Alexius 65, 243
- mental entities 11, 41–2, 68
- mention** *see* use
- metalanguage 84, 85, 176, 179–81, 189, 190
- metalinguistic variables 180
- metaphor 168–70, 172

- metaphysical**, metaphysics xvi, 2, 4, 8, 78, 83, 85, 89, 90, 119, 124, 126, 134, 153, 225
 methodological solipsism 86
 Mill, John Stuart 27, 127, 206, 208
 modality 123, 125, 148–50, 223
monadic predicates *see* one-place predicates
mood 8–9, 157–8, 160, 171, 190
 Moore, G. E. 235
Moore's paradox 235, 246
 Morning Star 29–31, 36, 47–8, 194
 Mozart 125–6, 138n5, 142
 multiple hyper-intensional embedding 197–8, 209
multiple relation theory of judgement 71–3
- naïve semantics** 3, 15–22, 23, 26–7, 29–33, 36, 47–8, 53–4, 65, 67n1, 70, 127
 narrow scope 53, 182, 199
 natural expressions 102
natural kinds 141–3, 153–4, 156n2, 205, 206, 225, 250n1, 117, 221
 natural language 105–6, 212–13
 naturalism 105–6, 118–19, 211–16
necessary *see* necessity
 necessity 101, 122–36, 142, 146–8, 198–9
negation 23, 56–7, 75, 83
negative existentials 37, 57–9, 167
 Neo-Meinongianism 243–4, 247
 Nessie *see* Loch Ness Monster
 Neurath, Otto 83, 85, 86–7, 91, 92, 106
 nonexistence 37, 58, 66, 167, 196–7, 244
non-extensional 40, 123, 179, 194–5, 198–9
 nonsense 42, 76–7, 83
 norms, **normative** 100, 223, 236–7, 246, 248, 260, 264
 'nothing' 84
- Oakington Puckshill Amber Sunblush 150
 Obama, Barack 133
- object language 84–5, 176, 177, 179–80, 189, 190, 158
 objects 16, 71, 76
observation sentences 108–9, 117, 120
 observation statements 87
occasion sentences 185, 191; *see also* standing sentences
 Octavian, Emperor 10
 Odysseus 242–5
 omniscient interpreter 187
one-place predicates 17, 19–21, 26, 27, 29, 32, 34, 48, 56, 68
 ontology 85, 105, 117, 262
 opacity, referential 39–41, 194–7, 200, 208
 open sentences 24–5, 37, 182
 open-textured 96
 ordinary language philosophy 167, 171
 other minds 80–1, 90, 104
- Paderewski, Ignacy Jan 205, 206
Paradox of Analysis 261, 272
 parameters 223–4, 228, 232
 paratactic 189, 208
 Partee, Barbara 153
 Peano, Giuseppe 46, 261
 Peregrin, Jaroslav 245
 performance *vs.* competence 216, 226–8
 performative utterances 158–60
perlocutionary acts 160, 162, 163, 170, 172, 173
 Perry, John 151, 153, 203
 perspective 255–6, 257, 265, 272
phenomenalism 79
phenomenological reductionism 79, 83
 Phocas, Emperor 11–12
picture theory 69, 72–5, 76, 77, 93, 84–5
 Plato 3, 26, 99, 178, 181–3, 215, 245, 259; 's problem 214–6
 pointing device 131
 politically correct 265
 positivism 85
possible 122–4, 135, 243, 247; *see also* necessary
 possible worlds 122, 124, 126, 127–9, 133–4, 135–6

- pragmatically imparted information 207
- pragmatically improper** 163–4, 207, 253
- pragmatics** xvii, 2, 3, 4, 9, 157–72, 174n3, 237, 245–6, 253, 266
- predicates** 16–21, 26, 193; Frege on 34, 35, 36, 39, 43–5; Quine on 115; Russell on 54, 60, 63–4, 59, 68–9, 70–1
- predication 5
- prescriptive 268
- presupposition** 56, 62, 161, 166–8, 172, 230, 253, 254, 272
- pretense theory 244
- principle of acquaintance 64, 66, 102, 115
- principle of charity** 186–7, 191
- Principle of Tolerance 84, 90
- private language** 19–20, 87, 101–2
- producers 129, 131
- pronouns 16, 25, 139, 140, 151–2, 203, 222, 223; *see also* anaphoric; indexicals
- proper names: Chomsky on 225; Frege on 31; Kripke on 124–9, 132, 135–6, 204–6; Quine on 115; Russell on 56, 60–1, 66; Salmon on 206–8; *see also* logically proper names; singular terms
- properties 21–2, 23, 27, 60, 63, 69, 71, 115
- propositional attitudes** xvi, 4, 11–12, 30, 36, 39–41, 53, 62, 66, 118, 190, 194–8, 199–210, 255; *see also* belief
- propositional content 47, 113, 120, 253
- propositional signs 73–4, 91
- propositions** 8–12, 27, 157–8, 256–7; Ayer's Verification Principle 78–9, 90; Frege on 30, 34, 39–41, 42, 197–8; Frege-Geach problem 82; Russell on 54, 64, 69–71, 89; *see also* atomic propositions; descriptivist theory of proper names; propositional attitudes; rigid designation
- 'Protocol Debate' 85–7, 88, 91
- pure indexicals** 140, 150; *see also* indexicals
- pure predicates** 18–19, 27
- Putnam, Hilary xvi, 4, 135, 141–4, 148, 153, 154, 198, 205, 225, 250n1, 257, 270, 272
- qualia 102
- quantifiers** 24–6, 75–6, 142, 200, 218; context-relativity 238, 247; Russell on 55, 125
- quietism 100
- Quine, W. V. xvi, 4, 5, 65, 79, 85, 89, 60–1, 84, 105–20, 135; and Chomsky 224, 225; and Davidson 175, 183, 190, 154–5; on exportation 199–201, 208, 209; holism of 113
- quotation, quotation marks 6, 194–5, 196, 211n1, 252
- racism 252–4, 257
- radical contextualism 241–2
- radical empiricism 81
- radical interpretation** 183–5, 191
- radical translation** 108–11, 118, 120
- Ramsey, Frank 77
- Rawls, John 235
- real Americans 265
- recursion** 13, 177, 178, 182, 229
- reductionism** *see* phenomenological reductionism
- reference** xvi, 3, 4, 5, 15–18, 21, 27, 216, 224; Chomsky on 225–6; context-relativity 139, 237–8; Davidson on 175, 181–2, 191; direct reference theory 127–32; fictional objects 242–5, 247; Frege on 29–47, 125; hyper-intensional embedding 197–8, 209; Kripke on 129–31; natural kind terms 142; opacity and transparency 194–5, 201; of predicates 43–5, 73–4; presupposition 166–8; Quine on 109–11, 116–18; Wittgenstein on 93–5, 104
- referential use** 167–8
- referentially opaque** 194–6, 200, 201, 209
- referentially transparent** 194–5, 200
- referents 16, 20–2, 27, 29–33
- reflective equilibrium 268–9, 273
- reflexive pronouns 140

- regulative rules 235–6, 246
- Reichenbach, Hans 88
- relational predicates 34
- relations 21–2, 27, 46, 60, 70, 71, 115, 245, 262
- relativism, epistemic 240, 247
- relevance theory xvi, 171
- replacement 114–16, 120, 262–4
- response and disposition 100
- Richard, Mark 208
- rigid designation 127–9, 136, 141–2, 144–6, 148–50, 153–4, 259
- Rorty, Richard 190
- rules 42, 130, 147; of assertion 235–7; in Chomsky 215, 220, 221, 223–4, 228; of conversation 163–4, 172; Wittgenstein on 94, 98–102
- Russell, Bertrand xv, xvi, 3, 27, 52–66, 212, 215, 219, 253, 270; and Frege 46–7; Logical Atomism 80; Multiple Relation Theory of judgement 64–8, 84; theory of definite descriptions 54–61, 124–5, 134, 166–8, 206, 242–3; and Wittgenstein 88–9, 102, 103
- Russell's paradox 68–9
- Ryle, Gilbert 171
- Salmon, Nathan 206–9, 210
- satisfy** 20, 34, 39, 180–1; *see also* predicates; properties
- saturation 45, 48, 49
- Saussure, Ferdinand de 216
- scare quotes 6, 89
- Schiffer, Stephen xvi, 171
- Schlick, Morris 78, 82, 85, 87
- Schoenberg, Arnold 85
- science of language 105–6, 212, 214–16, 224–8
- scope (wide, narrow)** 57, 182, 196, 199–202, 208, 210, 242, 244, 247
- Searle, John xv, xvi, 4, 157, 160–1, 171, 226, 235, 242
- second-order 37, 44, 116
- Sellars, Willard xv, 245
- semantic externalism 143–4, 153, 257–8
- semantics** 2, 160–1, 175–89, 224–8; meaning and reference 168, 172; naïve semantics 15–27, 29–32, 70; T-sentences 180–2
- sense**: Davidson on 178; Frege's theory of 32–7, 41–2, 45–6, 48–9, 151–2; hyper-intensional embedding 197–8, 209; Quine on 113; slurs 251–2, 253; *see also* descriptivist theory of proper names; propositional attitudes
- sense-contents 75–7
- sense-data** 42, 62–4, 66, 79–80, 102, 132,
- sentence-connectives** 23–4, 75, 123, 194
- sentence-meaning 146
- sentences 33–6, 85–7, 109, 185; open and closed 182–3
- sets 68–9, 70, 262–3
- shortest spy, the 202–3, 209
- significance, criterion of 78
- simple singular terms 31, 54, 191
- singular propositions** 64, 202
- singular terms** 11, 16–17, 19, 27, 29, 30, 32–6, 37, 40, 43, 45, 47–8, 52–4, 50, 55, 56–64; *see also* definite descriptions; indexicals; proper names
- slurs** 251–8
- Snoopy 245–6, 247
- Soames, Scott 208
- social context 100, 266–71 *passim*
- Socrates 26, 34, 71, 99, 115, 178, 181, 215, 260, 261
- speaker meaning and reference 167–8, 172
- speech-acts** xvi, 7, 170, 172, 173, 192, 234, 237, 252–3, 255; speech-act theory 158–62
- Sperber, Dan 171
- Stalnaker, Robert xvi, 134
- standing sentences** 108–9, 185, 191; *see also* occasion sentences
- state of affairs *see* facts
- statements of value 76–7, 81–3, 251
- Stoics 29
- Strawson, Peter F. 112, 166–8, 171, 212, 225, 226
- structure 12, 70
- subjective sentences 108
- subjectivism 240

- subjunctive conditionals *see*
 - counterfactual conditionals
- substitutivity** 38–40, 51, 53, 193–4; *see also* extensionality
- surface form, S-form, S-structure 56, 220–2, 233n1
- symbolic language 51n1, 212
- synonymy 9, 19, 61, 109, 112, 118, 125, 131, 133, 134, 136, 142, 189, 204, 210, 263
- syntax** 2, 3, 17–19, 23, 35, 83, 84, 171, 180–1, 183, 212ff
- synthetic truths 7, 82, 85, 89, 112, 126, 134

- tacit indexicals 205, 238, 250
- Tarski, Alfred xv, 23, 88, 135, 180, 190
- tautologies 76, 78, 90
- temporality 109–10, 148–50, 154
- tense** 10, 140, 145, 151, 187, 223
- that-clauses 11, 8, 40–1, 188–9
- theories of meaning 1–5, 15, 21, 132–4;
 - Davidson on 175–80, 183–7; Quine on 105–12, 118–19
- thick concepts** 251, 260, 264, 274
- thoughts 1, 13, 42, 45, 152, 228; *see also* propositions
- tokens** 5, 12
- tone 253–4, 265
- tool analogy 97–8, 168
- traces 220–1
- training *see* dispositions
- transformational grammar 219–21
- translation 108–11, 117–18, 119, 120, 178, 185, 211, 226
- transparency, referential 194–5, 200
- Travis, Charles 242, 271
- truth 21–2, 69–72, 75, 158–9, 206
- truth-conditions** 2, 6, 7, 15, 51, 179–83
- truth-functional** 23, 27, 68, 123, 124, 240–1; implicature 164–6; slurs 254
- truth-functional connectives *see*
 - conditionality; conjunctions;
 - disjunctions; negation
- truth-tables 75, 124, 240, 245
- truth-value 7, 193, 240; Frege on 35, 38–9, 36, 43

- T-sentences** 178–82, 191
- ‘Twin Earth’ argument (Putnam) 141–6
- two-dimensional semantics 146–8, 154
- two-place predicates** 17, 19, 116, 182, 192n2, 196, 261
- types** 5, 73

- understanding 12–14, 178; and sense
 - 34–5, 41–2, 105; Wittgenstein on 93–4, 100, 103
- universal grammar 28n2, 213, 215–23
- universal quantifiers** 24–5, 182–3
- universals** 21, 27, 60, 62–4, 65, 71, 76, 259
- use** 2, 5, 97, 157ff
- utterance *see* context of utterance

- value 77–8, 81–2, 90, 239, 251, 264, 265–8; *see also* truth-value
- Van Fraassen, Bas 82
- variables** 24–5, 182–3, 188–9
- verbs 18–19, 143, 159–60, 198, 217
- verification principle** 78–85, 90
- Vienna circle 82, 83, 85, 88
- vivid designators 202–4, 209
- Vlach, Frank 153
- Vulcan 31, 36–7, 45, 48, 51, 52–3, 132, 196–7, 200, 209

- Waismann, Friedrich 83, 85, 88
- Whitehead, Alfred North 47, 65, 69
- wide scope 57
- Williamson, Timothy 236, 271
- Wilson, Deirdre 171
- Wittgenstein, Ludwig iii, xvi, 1, 3, 5, 27, 46–7, 65, 69, 170, 216, 226, 228, 230, 242, 271; *Philosophical Investigations* 93–103; *Picture Theory* 72–7; *Tractatus Logico-Philosophicus* 27, 65, 69, 68–73, 88–90, 93, 94
- woke 265
- woman, women 266–8
- Wonder, Stevie 5
- Wright, Crispin 78, 102

- x-bar theory 221
- X-Phi 268–71, 273



Taylor & Francis Group
an informa business

Taylor & Francis eBooks

www.taylorfrancis.com

A single destination for eBooks from Taylor & Francis with increased functionality and an improved user experience to meet the needs of our customers.

90,000+ eBooks of award-winning academic content in Humanities, Social Science, Science, Technology, Engineering, and Medical written by a global network of editors and authors.

TAYLOR & FRANCIS EBOOKS OFFERS:

A streamlined experience for our library customers

A single point of discovery for all of our eBook content

Improved search and discovery of content at both book and chapter level

REQUEST A FREE TRIAL

support@taylorfrancis.com



Routledge
Taylor & Francis Group



CRC Press
Taylor & Francis Group